

# La convenzione

Studio filosofico

D. K. Lewis

**The Robin Hood Online Press**



**per Steffi**

**Traduzione italiana di Gabriele Usberti**



# Indice

<b>Ringraziamenti</b>	<b>iii</b>
<b>Prefazione</b>	<b>v</b>
<b>1 Coordinazione e convenzione</b>	<b>5</b>
1.1 Problemi di coordinazione campione . . . . .	5
1.2 Analisi dei problemi di coordinazione . . . . .	8
1.3 Risoluzione dei problemi di coordinazione . . . . .	23
1.4 Convenzione . . . . .	33
1.5 Convenzioni campione . . . . .	39
<b>2 Raffinamento della convenzione</b>	<b>49</b>
2.1 Conoscenza comune . . . . .	49
2.2 Conoscenza delle convenzioni . . . . .	57
2.3 Alternative alle convenzioni . . . . .	63
2.4 Gradi di convenzione . . . . .	70
2.5 Conseguenze delle convenzioni . . . . .	74
<b>3 Confronto tra convenzione e altre nozioni</b>	<b>77</b>
3.1 Accordo . . . . .	77
3.2 Contratti sociali . . . . .	82
3.3 Norme . . . . .	90
3.4 Regole . . . . .	93
3.5 Comportamento conformativo . . . . .	99
3.6 Imitazione . . . . .	109
<b>4 Convenzione e comunicazione</b>	<b>113</b>
4.1 Segnali campione . . . . .	113
4.2 Analisi della segnalazione . . . . .	120
4.3 Segnalazione verbale . . . . .	132
4.4 Significato convenzionale dei segnali . . . . .	133
4.5 Significato <sub>nn</sub> convenzionale dei segnali . . . . .	141

<b>5</b>	<b>Convenzioni di linguaggio</b>	<b>149</b>
5.1	Lingue possibili . . . . .	149
5.2	Grammatiche . . . . .	154
5.3	Semantica in una lingua possibile . . . . .	160
5.4	Convenzioni di veracità . . . . .	164
5.5	Semantica in una popolazione . . . . .	179
	<b>Conclusioni</b>	<b>187</b>

# Ringraziamenti

Ho un gradito debito con Willard Van Orman Quine per il suo incoraggiamento e il suo aiuto nel corso del mio tentativo di riabilitazione dell'analiticità. Sono riconoscente ai miei amici per i loro commenti alla prima e alla successiva versione. Un ringraziamento particolare va a George Boolos, Charles Chastain, David Kaplan, Ewart Lewis, Stephanie Lewis, Barbara Partee, Stephen Schiffer, Michael Slote, J. Howard Sobel, e agli studenti del mio seminario sulla filosofia del linguaggio tenuto all'University of California, Los Angeles nell'ultimo trimestre del 1967. Un vivo ringraziamento d'altro genere va alla C'est Si Bon Pâtisserie, dove è stata scritta la maggior parte di questo volume, per la sua ospitalità e il suo buon caffè. Ringrazio David Shwayder per avermi permesso di fare lunghe citazioni dal suo libro *The Stratification of Behavior*.

D. K. L.

Los Angeles  
Giugno 1968



# Prefazione

Quando ero bambino mi figuravo che il linguaggio fosse fissato e vagliato da un collegio di alti funzionari, riuniti in solenne assemblea attorno a un tavolo nello stile di Rembrandt. Per un certo tempo questa rappresentazione non fu turbata dal problema di sapere quale lingua avrebbero potuto usare i funzionari nelle loro deliberazioni, o dalla paura di un regresso infinito.

Immagino che questa raffigurazione sia stata condivisa da molti, nell'infanzia priva di senso critico. Certo, molti pensatori maturi hanno chiamato convenzionale il linguaggio. Molti, anche in altri campi, si sono facilmente richiamati ad accordi che, storicamente, non sono mai stati stipulati. Il contratto sociale, nella teoria dello Stato di Hobbes, ne è l'esempio più vistoso. Questo esempio è logicamente più rispettabile di quello del linguaggio, giacché la dottrina secondo cui lo Stato ebbe letteralmente origine con un contratto sociale non comporta un regresso infinito.

Non che i sostenitori della dottrina del contratto sociale intendano essere interpretati così alla lettera, naturalmente; essi vogliono dire soltanto che lo Stato è *come se* fosse stato costituito in questo modo. Ma, allora, questo "come se" pone il problema, in termini psicoanalitici, del suo contenuto latente: in che modo, esattamente, lo Stato assomiglia a ciò che un vero e proprio contratto sociale avrebbe potuto darci? Nel caso del linguaggio il problema del contenuto latente è ancora più grave e imbarazzante, in quanto una fondazione originaria del linguaggio per pubblica convenzione non è soltanto storicamente falsa ma anche inconcepibile. Dov'è la convenzione, se non può esserci neppure l'idea di stipulare socialmente una convenzione?<sup>1</sup>

Alcuni filosofi hanno dato grande peso a una pretesa distinzione tra verità *analitiche*, vere esclusivamente per convenzione linguistica, e verità *sintetiche*, che dicono qualcosa di rilevante intorno alla realtà. Questo è un caso caratteristico e decisivo di appello alla convenzione laddove non avrebbe potuto esserci l'idea di stipulare socialmente una convenzione. Infatti, i filosofi

---

<sup>1</sup>Gioco di parole tra *convention* (convenzione) e *convene* (riunirsi, radunarsi). (N.d.T.)

in questione considerano analitiche le verità logiche; e qui sorgerebbe una circolarità, se dovessimo formulare le convenzioni in modo esplicito. La difficoltà sta nel fatto che, per applicare una qualunque convenzione esplicita abbastanza generale da produrre tutte le verità logiche, dovremmo già far uso della logica nel trarre inferenze dalle convenzioni generali per giungere alle applicazioni individuali.

Abbiamo davanti a noi uno studio, insieme lucido e fantasioso, divertente e meticoloso, in cui Lewis si propone di rendere la nozione di convenzione indipendente da ogni reale o immaginata stipulazione sociale. Egli si propone di isolare i tratti peculiari della convenzionalità, il contenuto latente, senza giovare di analogie e di finzioni. Molto approssimativamente, la nota dominante della convenzionalità è una certa indifferenza: per quanto ci riguarda, la sillaba “sì” avrebbe potuto significare “no”, la luce rossa avrebbe potuto significare “avanti”, e le cravatte nere avrebbero potuto esser considerate meno formali di quelle fantasia. Questa è l’intuizione iniziale; ma il senso appropriato dell’indifferenza, o di “avrebbe potuto significare”, necessita di un notevole raffinamento. Lo scopo è raggiunto, grazie all’abile uso che Lewis fa della più recente teoria dei giochi e della decisione.

Il problema di distinguere tra verità analitiche e sintetiche era evidentemente uno dei moventi della ricerca. Alla fine, Lewis conclude che la nozione di convenzione non è il punto cruciale della distinzione. Ma ciò non comporta che egli trovi inaccettabile la nozione di analiticità. Al contrario, egli si chiede perché io abbia messo in dubbio principalmente questa nozione, più ancora di quella di verità. Non è questa la sede per ripetere i motivi per cui secondo me la verità è la nozione di gran lunga più solida;<sup>2</sup> ad ogni modo, egli deve aver trovato le mie ragioni non convincenti. Piuttosto, egli si ferma dove altri hanno cominciato, fondando la nozione di analiticità su quella di mondo possibile. Il fatto che egli si accontenti di questa impostazione del problema dell’analiticità ci induce a chiederci, in sostanza, quale abbia potuto essere la motivazione fondamentale del suo studio sulla convenzione; ma, qualunque sia stato il movente, dobbiamo essergli riconoscenti, perché nel corso del libro il lettore arriva a scoprire nella convenzione, e non nell’analiticità, un concetto-chiave della filosofia del linguaggio.

W.V.O. Quine

Harvard University  
26 settembre 1968

---

<sup>2</sup>Cfr. *From a Logical Point of View*, p. 138, ed eventualmente anche *The Ways of Paradox and Other Essays*, p. 123. [Del primo testo citato esiste una traduzione italiana di E. Mistretta: *Il problema del significato*, Ubaldini, Roma, 1966. (N.d.T.)]

# Introduzione

La professione del filosofo consiste nel mettere in dubbio i luoghi comuni che altri accettano senza pensarci due volte. Una professione pericolosa, perché è più facile dubitare di un filosofo che di un luogo comune che sia utile. Infatti, quando un buon filosofo mette in dubbio un luogo comune, di solito risulta che il luogo comune era sostanzialmente giusto; ma il filosofo ha notato delle difficoltà che una persona che non ci pensa due volte non avrebbe potuto incontrare. Alla fine il dubbio riceve una risposta e, nella maggior parte dei casi, il luogo comune sopravvive. Ma il filosofo ha reso un servizio ai sostenitori del luogo comune: li ha indotti a pensarci due volte.

È un luogo comune che il linguaggio sia regolato da convenzioni. Le parole potrebbero essere usate per significare pressoché tutto; e noi che le usiamo abbiamo dato loro il significato che hanno perché in qualche modo, per gradi e intuitivamente, siamo giunti alla consapevolezza che è con questo significato che le useremo. Potremmo benissimo usare queste parole in modo diverso - o usare parole diverse, come fanno le persone di paesi stranieri. Volendo, potremmo cambiare le nostre convenzioni.

Dire soltanto questo non è molto. Non equivale a descrivere il linguaggio come un calcolo, preciso e rigoroso, né a contrapporre il discorso “corretto” a quello colloquiale, o viceversa. Non equivale a dire che tutte le lingue che possiamo concepire sono ugualmente buone, o che ogni aspetto di una lingua efficiente avrebbe potuto essere diverso. Non equivale a dire che le verità necessarie sono create per convenzione, ma solo che le verità necessarie, come le verità geologiche, sono espresse convenzionalmente con queste parole piuttosto che con quelle. Non equivale a esaltare i poteri della convenzione come fanno alcuni filosofi “convenzionalisti”, ma solo ad insistere sul fatto che la convenzione esiste. Il luogo comune secondo il quale ci sono convenzioni di linguaggio non è il dogma di alcuna scuola filosofica, ma impone l’assenso immediato a ogni persona capace di riflettere - a meno che sia un filosofo.

Questo semplice luogo comune, infatti, è stato posto sotto accusa. W.V.O.

Quine lo mise in dubbio nel 1936 e in seguito lo ripudiò completamente.<sup>1</sup> Morton White si unì a lui nella lotta<sup>2</sup> e, insieme, essi persuasero molti a condividere i loro dubbi, e molti di più ridussero al silenzio. Quine e White argomentano che le supposte convenzioni di linguaggio non possono essere molto simili ai casi tipici e chiari di convenzione. Le convenzioni sono accordi; ma noi ci siamo mai accordati reciprocamente per conformarci, nel nostro uso del linguaggio, a regole stipulate? No. Se l'hanno fatto i nostri antenati, perché questo dovrebbe riguardare noi, che ce ne siamo dimenticati? In ogni caso, le convenzioni di linguaggio non avrebbero eventualmente potuto trarre origine da un accordo, perché alcune di esse sarebbero occorse per fornire quella lingua rudimentale in cui fu stipulato il primo accordo. Noi non possiamo neppure dire quali siano le nostre convenzioni, se non con un lungo procedimento di prova ed errore. Le conoscevamo meglio quando le adottammo per la prima volta? Noi non disponiamo di un concetto di convenzione che permetta al linguaggio di essere convenzionale; siamo inclini a chiamare convenzionali alcuni aspetti del linguaggio, ma non sappiamo dire perché. Possiamo cedere a questa inclinazione Quine stesso lo fa<sup>3</sup> - ma ciò non aumenta affatto la nostra comprensione del linguaggio. Conclusione: le convenzioni linguistiche sono un mito. La nuda verità è che il nostro uso del linguaggio si conforma a certe regolarità; e questo è tutto.

Noi possiamo controbattere disperatamente che ci deve essere *qualcosa* di valido nella nostra nozione di convenzione di linguaggio, anche se non sappiamo dire cosa. Quando veniamo a contatto con tale nozione noi tutti *riusciamo* a padroneggiarne l'idea, e ci comportiamo tutti più o meno allo stesso modo nel distinguere tra aspetti del linguaggio che chiamiamo convenzionali e aspetti non convenzionali. Quindi bisogna che tale nozione abbia un senso. La convenzionalità dev'essere perlomeno quella cosa (non sappia-

---

<sup>1</sup>“Truth by Convention”, *Philosophical Essays for A.N. Whitehead*, a cura di O.H. Lee, Longmans, New York, 1936 “Two Dogmas of Empiricism”, *From a Logical Point of View: Nine Logico-Philosophical Essays*, Harvard University Press, Cambridge, Mass., 1961, II edizione, pp. 20-46; [Trad. it. cit., pp. 20-44 (N.d.T.)] “Carnap and Logical Truth”, *The Philosophy of Rudolf Carnap*, a cura di P.A. Schlipp, Open Court, La Salle, 1963, pp. 385-406; *Word and Object*, MIT Press, Cambridge, Mass., e John Wiley, New York, 1960. [Trad. it. di F. Mondadori, *Parola e oggetto*, Il Saggiatore, Milano, 1970. (N.d.T.)]

<sup>2</sup>“The Analytic and the Synthetic: An Untenable Dualism”, *John Dewey: Philosopher of Science and Freedom*, ed. Sidney Hook, Dial, New York, 1950, pp. 316-330; [Trad. it. di A. Meotti, “Analitico e sintetico: un dualismo insostenibile, in *Semantica e filosofia del linguaggio*, a cura di L. Linsky, Il Saggiatore, Milano, 1969, pp. 339-354 (N.d.T.)] *Toward Reunion in Philosophy*, Harvard University Press, Cambridge, Mass., 1956.

<sup>3</sup>Alla fine di “Carnap and Logical Truth”, dove dice: “La cultura dei nostri padri è una cultura grigio-pallida, nera di fatti e bianca di convenzioni.”

mo quale) che evoca una risposta caratteristica in chiunque abbia ricevuto il nostro stesso tipo di educazione.

Ma come sarebbe meglio conoscere ciò di cui stiamo parlando! Ossia disporre di un'analisi della convenzione di assoluta generalità, tale da comprendere la convenzione tacita, non istituita per accordo. Questo libro è il mio tentativo di compiere un'analisi di questo tipo. Spero che sia un'analisi del nostro comune e consolidato concetto di convenzione, in modo che voi riconosciate che spiega quello che dovevate avere in mente quando dicevate che il linguaggio - come molte altre attività - è regolato da convenzioni. Ma forse non lo è, perché forse non tutti condividiamo un chiaro concetto generale di convenzione. Per lo meno, nella misura in cui *io* avevo un concetto di convenzione prima di pensarci due volte, quello che discuterò è o quel concetto o il suo legittimo erede. E ciò che io chiamo convenzione è un fenomeno importante con qualunque nome lo si chiami. Il linguaggio è soltanto una delle numerose attività regolate da convenzioni che non abbiamo stipulato per accordo e che non siamo in grado di descrivere.

La mia teoria della convenzione ha avuto origine dalla teoria dei giochi di coordinazione pura - una branca trascurata della teoria generale dei giochi di von Neumann e Morgenstern, molto diversa dal punto di vista del metodo e del contenuto dalla loro fortunata e più nota teoria dei giochi di conflitto puro. I giochi di coordinazione sono stati studiati da Thomas C. Schelling,<sup>4</sup> ed è stato lui a fornirmi l'occorrente per costruire una risposta a Quine e White.

In fondo, però, la teoria dei giochi non è che l'impalcatura. Posso riformulare la mia analisi della convenzione senza farne uso. Il risultato è una teoria che si ispira a quella elaborata da Hume nella sua discussione sull'origine della giustizia e della proprietà. La convenzione si rivela

una consapevolezza generale per l'interesse comune, consapevolezza che tutti i membri della società esprimono l'un l'altro, e che li induce a regolare la loro condotta in base a certe regole. Osservo che è nel mio interesse lasciare a un altro il possesso dei suoi beni, purché egli agisca nello stesso modo nei miei confronti. Anche l'altro è consapevole di un analogo interesse a regolare la sua condotta. Quando ci si esprime reciprocamente questa consapevolezza dell'interesse comune, allora essa produce una risoluzione e un comportamento adeguati. E questo, di certo, si può chiamare abbastanza propriamente una convenzione o un accordo tra di noi, anche se manca qualsiasi promessa, dato che

---

<sup>4</sup>*The Strategy of Conflict*, Harvard University Press, Cambridge, Mass., 1960.

le azioni di ciascuno di noi sono in rapporto con quelle altrui e le compiamo in base alla supposizione che l'altro dovrà compierne certe altre.<sup>5</sup>

---

<sup>5</sup>A *Treatise of Human Nature*, III.ii.2. [Trad. it. di E. Lecaldano e E. Mistretta, *Trattato sulla natura umana*, III, II, 2., in *Opere*, Laterza, Bari, 1971, Vol. I. (N.d.T.)]

# Capitolo 1

## Coordinazione e convenzione

### 1.1 Problemi di coordinazione campione

L'uso del linguaggio appartiene a una classe di situazioni con un importante carattere in comune: situazioni che chiamerò *problemi di coordinazione*. Non ne darò una definizione finché non avremo visto alcuni esempi. Cominciamo con situazioni che si potrebbero presentare tra due persone - persone che chiameremo "tu" e "io".

(1) Supponiamo che tu e io vogliamo incontrarci. Ci incontreremo se e solo se ci recheremo nello stesso posto. A ciascuno di noi due importa poco dove va (entro certi limiti), se poi vi incontra l'altro; parimenti, gli importa poco dove va, se poi non vi incontra l'altro. Ciascuno di noi due deve scegliere dove andare. Per me, il posto migliore in cui andare è quello in cui andrai tu, e per questo cerco di immaginare dove andrai tu per andarci anch'io. Tu fai lo stesso. Ciascuno sceglie in base a quella che si aspetta sia la scelta dell'altro. Se uno ha successo ce l'ha anche l'altro; l'esito da noi desiderato è lo stesso.

(2) Supponiamo che tu e io stiamo parlando al telefono e siamo inaspettatamente interrotti dopo tre minuti. Entrambi vogliamo che la comunicazione sia ristabilita immediatamente, il che avverrà se e solo se uno di noi richiama e l'altro attende. A ciascuno di noi due importa poco se egli sarà quello che richiama o quello che attende. Entrambi dobbiamo scegliere se richiamare, ciascuno in base a quella che si aspetta sia la scelta dell'altro, in modo da richiamare se e solo se l'altro attende.

(3) Un esempio tratto dal *Treatise of Human Nature* di Hume: supponiamo che tu e io stiamo remando insieme su una barca. Se vogliamo con ritmo la barca fila liscia e dritta; altrimenti la barca va piano e devia, sprechiamo

energie e rischiamo di incagliarci. Noi scegliamo continuamente se remare più in fretta o più adagio; ci importa poco con quale velocità remiamo, purché vogliamo con ritmo. Così ciascuno dei due regola continuamente la sua velocità per accordarla a quella che si aspetta che l'altro mantenga.

Passiamo adesso a situazioni con più di due persone.

(4) Supponiamo che parecchi di noi stiano guidando sulle stesse strade a due corsie e con numerose curve. A chiunque di noi importa poco di guidare sulla corsia destra o su quella sinistra, purché gli altri facciano altrettanto. Ma se qualcuno guida a destra e qualcuno a sinistra, ciascuno corre il pericolo di una collisione. Ognuno deve quindi scegliere se guidare a destra o a sinistra, in base a quello che si aspetta dagli altri: guidare a sinistra se la maggior parte degli altri guida a sinistra, a destra se la maggior parte degli altri guida a destra (e guidare sulla corsia che gli pare se gli altri sono suddivisi più o meno equamente).

(5) Supponiamo di essere campeggiatori che vanno in cerca di legna da ardere. A chiunque di noi importa poco in quale direzione va, ma se due vanno nella stessa direzione è probabile che esplorino lo stesso terreno, cosicché quello che arriva dopo non trova legna. Ciascuno deve scegliere la direzione da seguire in base a quello che si aspetta dagli altri: una direzione diversa da quella di tutti gli altri.

(6) Supponiamo che parecchi di noi siano stati invitati a un party. A chiunque importa poco del proprio abito. Ma egli si troverebbe a disagio se gli altri fossero vestiti in modo simile ed egli in modo diverso, perché egli sa che chi vuole può dare una spiegazione disonorevole di quella differenza. Ciascuno deve quindi vestirsi in base a come si aspetta che si vestiranno gli altri: con lo smoking se gli altri indosseranno lo smoking, con un completo da clown se gli altri indosseranno un completo da clown (e come gli pare se gli altri saranno vestiti in fogge diverse).

(7) Supponiamo di essere degli ogopolisti scontenti. Al varare del prezzo della nostra materia prima ciascuno di noi deve stabilire nuovi prezzi. Non è vantaggioso per nessuno stabilire dei prezzi più alti di quelli fissati dagli altri, perché se uno lo fa rischia di perdere la sua fetta di mercato. Né è vantaggioso per alcuno stabilire dei prezzi più bassi di quelli fissati dagli altri, perché se uno lo fa minaccia i suoi concorrenti e incorre nelle loro rappresaglie. Ciascuno deve quindi fissare i suoi prezzi entro la gamma dei prezzi che si aspetta siano stabiliti dagli altri.

(8) Un esempio tratto dal *Discours sur l'inégalité* di Rousseau: supponiamo di trovarci senza cibo in un luogo deserto. Da soli possiamo catturare conigli e mangiar male. Insieme possiamo catturare cervi e mangiar bene. Ma se uno soltanto di noi deserterà la caccia al cervo per catturare un co-

niglio, il cervo scapperà via e gli altri cacciatori non mangeranno se non disenteranno anch'essi. Ciascuno deve scegliere se partecipare o no alla caccia al cervo in base a quello che si aspetta dagli altri, partecipando se e solo se nessun altro disenterà.

(9) Supponiamo di ritenere nostro interesse comune dividere un magro avere, per esempio un pascolo, in parti tali che ciascuno di noi possa fare assegnamento sullo sfruttamento esclusivo di una porzione di esso. (Supponiamo che nessuno pensi neppure che sarebbe suo interesse approfittare della porzione di un altro. Lo sforzo, il torto inflitto al vicino, il cattivo esempio, la generale perdita di fiducia sembrano invariabilmente più importanti di qualsiasi guadagno.) A chiunque importa poco chi sfrutta una certa porzione e qual è questa porzione, purché più persone non cerchino di sfruttare la stessa porzione e nessuna porzione vada sprecata. Ciascuno deve scegliere quale porzione sfruttare in base alle porzioni che si aspetta che gli altri sfrutteranno e a quella che gli lasceranno.

(10) Supponiamo di essere commercianti. A chiunque di noi importa poco quali merci (diverse da quelle che egli stesso può usare) riceve in cambio delle sue. Ma se egli accetta ciò che gli altri rifiutano si trova ingombro di qualcosa di inutile; se invece rifiuta ciò che gli altri accettano reca inutilmente disturbo ai suoi clienti e a se stesso. Ciascuno deve scegliere quello che accetterà in base a ciò che si aspetta di poter spendere - cioè in base a ciò che gli altri accetteranno: oro e argento se può spendere oro e argento, banconote degli Stati Uniti se può spendere banconote degli Stati Uniti, denaro canadese se può spendere denaro canadese, collane di conchiglie se può spendere collane di conchiglie, capre se può spendere capre, qualunque cosa capiti se può spendere qualunque cosa capiti, nulla se non può spendere nulla.

(11) Supponiamo di poter adottare, con la pratica, una qualunque lingua di una gamma piuttosto estesa. A chiunque importa relativamente poco (in una prospettiva a lunga scadenza) quale lingua adotta, purché egli e coloro che gli stanno vicino adottino la stessa lingua e possano comunicare facilmente. Ciascuno può scegliere quale lingua adottare in base a quello che si aspetta relativamente alla lingua dei suoi vicini: l'inglese tra parlanti inglesi, il gallese tra parlanti gallesi, l'esperanto tra persone che parlano l'esperanto, e così via.

## 1.2 Analisi dei problemi di coordinazione

Sulla base di questi esempi, vediamo come si può descrivere la caratteristica comune dei problemi di coordinazione.

Due o più agenti devono scegliere ciascuno una tra più azioni alternative. Spesso tutti gli agenti hanno lo stesso insieme di azioni alternative, ma ciò non è necessario. Gli esiti che gli agenti vogliono produrre o evitare sono determinati collettivamente dalle azioni di tutti gli agenti. Così l'esito di qualunque azione un agente possa scegliere dipende dalle azioni degli altri agenti. È per questo motivo - come abbiamo visto in ogni esempio - che ciascuno deve scegliere che cosa fare in base a quello che si aspetta che faranno gli altri.

Alcune combinazioni delle azioni scelte dagli agenti sono *equilibri*: combinazioni in cui ciascun agente, date le azioni degli altri agenti, si è comportato nel modo migliore. In un punto di equilibrio nessun agente avrebbe potuto produrre un esito a lui più gradito comportandosi diversamente, a meno che alcune azioni degli altri fossero state diverse. Nessuno si pente della propria scelta quando viene a conoscenza delle scelte degli altri. Nessuno ha perso per mancanza di preveggenza.

Questo non equivale a dire che un punto di equilibrio debba produrre un esito che sia il migliore anche soltanto per uno degli agenti (anche se è vero che, se c'è una combinazione che è la migliore per ciascuno, tale combinazione deve essere un equilibrio). In un equilibrio alcuni o tutti gli agenti avrebbero potuto benissimo ottenere risultati migliori se alcuni o tutti si fossero comportati diversamente. Quello che non è possibile è che uno qualunque degli agenti ottenga risultati migliori comportandosi in modo diverso soltanto lui, mentre tutti gli altri si comportano nello stesso modo di prima.

Possiamo illustrare gli equilibri disegnando delle *matrici delle vincite* per i problemi di coordinazione tra due agenti. Chiameremo gli agenti *giocatore di riga* e *giocatore di colonna*. Rappresentiamo le azioni alternative del giocatore di riga con righe etichettate della matrice e quelle del giocatore di colonna con colonne etichettate. I quadrati rappresentano così le combinazioni delle azioni degli agenti, e quindi gli esiti attesi. I quadrati sono etichettati con due *vincite*, numeri che in qualche modo misurano la desiderabilità che l'esito atteso ha per il giocatore di riga e per quello di colonna.<sup>1</sup>

---

<sup>1</sup>La mia descrizione non richiederà una grande sottigliezza per quanto riguarda queste misure numeriche della desiderabilità. Una fondazione, se occorre, può essere fornita dalla teoria della decisione quale è stata sviluppata, per esempio, da Richard Jeffrey in *The Logic of Decision*, McGraw-Hill, New York, 1965. Penso che la teoria della decisione si ap-

La vincita del giocatore di riga è a sinistra in basso, quella del giocatore di colonna a destra in alto.

La matrice della Figura 1 potrebbe quindi rappresentare una versione semplice dell'esempio (1), dove R1, R2 e R3 sono le azioni del giocatore di

	C1	C2	C3
R1	1 si incontrano	0	0
R2	0	1 si incontrano	0
R3	0	0	1 si incontrano

Figura 1

riga consistenti nell'andare, rispettivamente, nei luoghi L1, L2 e L3, mentre C1, C2 e C3 sono le azioni del giocatore di colonna consistenti nell'andare, rispettivamente, nei luoghi L1, L2 e L3. Gli equilibri sono le tre combinazioni in cui entrambi i giocatori vanno nello stesso luogo e là si incontrano:  $\langle R1, C1 \rangle$ ,  $\langle R2, C2 \rangle$  e  $\langle R3, C3 \rangle$ . Per esempio,  $\langle R2, C2 \rangle$  è un equilibrio per definizione, in quanto il giocatore di riga lo preferisce a R1, C2 o a  $\langle R3, C2 \rangle$ , e il giocatore di colonna lo preferisce a  $\langle R2, C1 \rangle$  o a  $\langle R2, C3 \rangle$ . Entrambi non hanno preferenze fra i tre equilibri.

Ma supponiamo di modificare l'esempio in modo che ai giocatori importi dove vanno, anche se non tanto quanto gli importa incontrarsi. La nuova matrice delle vincite potrebbe essere quella della Figura 2. Gli equilibri rimangono gli stessi:  $\langle R1, C1 \rangle$ ,  $\langle R2, C2 \rangle$ , e  $\langle R3, C3 \rangle$ . Ma i giocatori non sono più indifferenti tra equilibri diversi.  $\langle R1, C1 \rangle$  è per entrambi il miglior esito possibile;  $\langle R3, C3 \rangle$  è per entrambi l'esito d'equilibrio peggiore, sebbene entrambi lo preferiscano ad esiti di non equilibrio. Se invece la matrice delle vincite fosse quella della Figura 3,  $\langle R1, C1 \rangle$  sarebbe per il

---

plichi, in modo approssimativo, a normali agenti razionali con preferenze imperfettamente coerenti; le nostre vincite non devono essere niente di più che indicazioni approssimative dell'intensità della preferenza.

	C1	C2	C3
R1	1,5 si incontrano 1,5	0,2 0,5	0 0,5
R2	0,5 0,2	1,2 si incontrano 1,2	0 0,2
R3	0,5 0	0,2 0	1 si incontrano 1

Figura 2

	C1	C2	C3
R1	1 si incontrano 1,5	0,2 0,5	0,5 0,5
R2	0 0,2	1,2 si incontrano 1,2	0,5 0,2
R3	0 0	0,2 0	1,5 si incontrano 1

Figura 3

giocatore di riga l'esito migliore e per il giocatore di colonna l'esito d'equilibrio peggiore;  $\langle R3, C3 \rangle$  sarebbe l'esito migliore per il giocatore di colonna e l'esito d'equilibrio peggiore per il giocatore di riga. Nessun esito sarebbe il migliore per entrambi.

Sembra che ci sia una differenza tra punti di equilibrio in cui ogni agente compie la stessa azione e punti di equilibrio in cui gli agenti compiono azioni diverse. Ma è una falsa differenza. Noi diciamo che gli agenti compiono la stessa azione se compiono azioni della stessa specie, azioni particolari che cadono sotto una descrizione comune. Ma le azioni possono essere descritte in un numero arbitrario di modi, nessuno dei quali ha un indiscusso diritto di priorità. Per *qualunque* combinazione di azioni, e *a fortiori* per qualunque punto di equilibrio di azioni, c'è *qualche* maniera di descrivere le azioni

alternative degli agenti in modo che cadano sotto una descrizione comune esattamente le azioni alternative della combinazione data. Qualunque combinazione, che sia un punto di equilibrio o no, è una combinazione di azioni di *una* stessa specie (una specie che esclude tutte le azioni alternative degli agenti). Se la si possa definire una combinazione in cui ogni agente compie la stessa azione dipende esclusivamente dalla naturalezza di quella classificazione.

Consideriamo l'esempio (2). Se abbiamo in mente queste descrizioni di azione:

R1 o C1: richiamare  
 R2 o C2: non richiamare

allora disegniamo la matrice delle vincite della Figura 4 e consideriamo l'esempio come un caso in cui gli equilibri  $\langle R1, C2 \rangle$  e  $\langle R2, C1 \rangle$  sono combinazioni in cui gli agenti compiono azioni diverse. Ma se abbiamo in mente queste descrizioni di azione:

R1' o C1': richiamare se e solo se si è colui che ha chiamato per primo  
 R2' o C2': richiamare se e solo se non si è colui che ha chiamato per primo

allora disegniamo la matrice delle vincite della Figura 5 e consideriamo l'esempio come un caso in cui gli equilibri  $\langle R1', C1' \rangle$  e  $\langle R2', C2' \rangle$  sono combinazioni in cui gli agenti compiono la stessa azione. Ma che cosa rende la prima coppia di descrizioni di azione più naturale della seconda? E se anche è così, che importa?

	C1	C2
R1	0	1
R2	1	0

Figura 4

	C1'	C2'
R1'	1	0
R2'	0	1

Figura 5

Potremmo dire che i problemi di coordinazione sono situazioni in cui più agenti cercano di conseguire un'uniformità di azione facendo ciascuno tutto quello che faranno gli altri. Ma esprimersi così è pericoloso, perché ciò che abbiamo detto è vero di un problema di coordinazione solo per descrizioni di azione adatte, e a volte le descrizioni che lo rendono vero ci danno un'impressione di artificiosità come negli esempi (2), (5), (9) e, forse, (4). Quello che è importante, a proposito delle combinazioni uniformi cui siamo interessati, non è che esse - per certe descrizioni - siano uniformi, ma che siano degli equilibri.

Naturalmente, questo non vuol dire che i problemi di coordinazione siano contraddistinti dalla presenza di equilibri. Anzi, la maggior parte della teoria matematica dei giochi è costituita precisamente dalla teoria dei punti di equilibrio (noti con il nome di *punti di sella* o *soluzioni*) in situazioni della specie opposta: conflitto puro di interesse tra due agenti, come nella Figura 6. In generale, il conflitto puro si può rappresentare per mezzo di una matrice delle vincite in cui le vincite degli agenti danno come somma zero in ogni quadrato (eventualmente dopo opportune trasformazioni lineari).<sup>2</sup> In altri termini, le perdite di un agente sono vincite per l'altro, e viceversa. Eppure, nel conflitto puro ci sono degli equilibri. Nell'esempio della Figura 6,  $\langle R1, C1 \rangle$  è un equilibrio: il giocatore di riga lo preferisce a  $\langle R2, C1 \rangle$  o a  $\langle R3, C1 \rangle$ , mentre il giocatore di colonna lo preferisce, a  $\langle R1, C2 \rangle$  o a

<sup>2</sup>Non c'è scopo a cambiare la definizione in modo da permettere che la somma sia una costante diversa da zero. Permettendo la trasformazione, abbiamo già raggiunto una generalità completa. Senza trasformazione, non raggiungeremmo una generalità completa semplicemente permettendo somme costanti diverse da zero. E permettendo la trasformazione lineare, rendiamo chiaro perché - nonostante le apparenze - la nostra definizione non dipende dal alcun problematico confronto interpersonale di desiderabilità.

	C1	C2	C3
R1	0	-0,5	-0,5
R2	0,5	1	-1
R3	0,5	-1	1

Figura 6

<R1, C3>.

Schelling sostiene l'opportunità di un "riorientamento della teoria dei giochi" in cui i giochi - cioè i problemi di decisione interdipendente - siano intesi variare su uno spettro avente come estremità opposte i giochi di conflitto puro e i giochi di coordinazione pura.<sup>3</sup> I *giochi di conflitto puro*, in cui gli interessi degli agenti sono perfettamente opposti, possono essere definiti come abbiamo appena visto. I *giochi di coordinazione pura*, in cui gli interessi degli agenti coincidono perfettamente, i sono giochi in cui le vincite degli agenti (eventualmente dopo opportune trasformazioni lineari) sono uguali in ogni quadrato. Gli altri giochi sono mescolanze, in proporzioni variabili, di conflitto e coordinazione, di opposizione e di coincidenza di interessi.

I miei problemi di coordinazione del tipo di (1)-(11) fanno parte delle situazioni che si trovano all'estremo di pura coordinazione dello spettro di Schelling (o vicino ad esso). Io non intendo richiedere una perfetta coincidenza di interessi. Per esempio, ho ammesso una coincidenza imperfetta di interessi nelle versioni dell'esempio (1), in cui ai giocatori importa alquanto del luogo in cui si recano, sebbene molto meno di quanto gli importa incontrarsi. Riprendiamo in considerazione le matrici delle vincite delle Figure 2 e 3. In parecchi quadrati le vincite non sono uguali. Nessuna trasformazione lineare di alcuna matrice potrebbe renderle uguali simultaneamente in ogni quadrato.

Intendo comunque limitare la mia attenzione alle situazioni in cui pre-

<sup>3</sup>*Strategy of Conflict*, pp. 83-118, 291-303.

domina la coincidenza di interessi: in cui, cioè, le differenze tra le vincite dei vari agenti in un qualunque quadrato (eventualmente dopo opportune trasformazioni lineari) sono piccole in confronto ad alcune delle differenze tra vincite appartenenti a quadrati diversi. È il caso delle matrici delle Figure 2 e 3; la differenza maggiore all'interno di un quadrato è di 0,5, mentre la differenza maggiore tra le vincite di quadrati diversi è di 1,5.

Un equilibrio, ricordiamo, è una combinazione in cui nessuno avrebbe conseguito esiti migliori se egli soltanto si fosse comportato diversamente. Definirò un *equilibrio di coordinazione* come una combinazione in cui nessuno avrebbe conseguito esiti migliori se un *qualunque* agente, sia lui che qualche altro, si fosse comportato da solo in modo diverso. In base alle definizioni, gli equilibri di coordinazione sono equilibri. Nei giochi di coordinazione pura gli equilibri sono sempre equilibri di coordinazione, in quanto gli interessi degli agenti coincidono perfettamente. Qualunque gioco di coordinazione pura ha almeno un equilibrio di coordinazione, in quanto ha almeno un esito che è il migliore per tutti. Ma gli equilibri di coordinazione non sono affatto limitati ai giochi di coordinazione pura. Sono comuni in situazioni in cui opposizione e coincidenza di interessi sono mescolate. Possono presentarsi anche in giochi di conflitto puro: nella Figura 7,  $\langle R1, C1 \rangle$  è un equilibrio di coordinazione.

	C1	C2
R1	0	0
R2	0	-1

Figura 7

La maggior parte delle versioni dei nostri problemi di coordinazione campione non sono giochi di coordinazione pura, ma hanno tutte degli equilibri di coordinazione. Abbiamo osservato che le versioni del problema del luogo di incontro illustrate nelle Figure 2 e 3 non sono giochi di coordinazione pura; ma i loro equilibri  $\langle R1, C1 \rangle$ ,  $\langle R2, C2 \rangle$  e  $\langle R3, C3 \rangle$  in entrambe le versioni sono equilibri di coordinazione.

Ciò non significa che *tutti* gli equilibri di un problema di coordinazione

debbano essere equilibri di coordinazione. Consideriamo una versione ulteriore dell'esempio (1). Supponiamo che ci sia una quarta località L4. Sia al giocatore di riga che a quello di colonna piace recarsi in L4, ma un incontro in L4 diminuirebbe il loro piacere di recarsi in L4, ed L4 sarebbe di poca utilità come luogo d'incontro. Abbiamo, così la matrice della Figura 8, con i soliti equilibri di coordinazione  $\langle R1, C1 \rangle$ ,  $\langle R2, C2 \rangle$ ,  $\langle R3, C3 \rangle$  e un nuovo equilibrio di non coordinazione,  $\langle R4, C4 \rangle$ . Si tratta di un equilibrio perché il giocatore di riga lo preferisce a  $\langle R1, C4 \rangle$ ,  $\langle R2, C4 \rangle$  o  $\langle R3, C4 \rangle$ , e il giocatore di colonna lo preferisce a  $\langle R4, C1 \rangle$ ,  $\langle R4, C2 \rangle$  o  $\langle R4, C3 \rangle$ . Non è un equilibrio di coordinazione perché non tutte - di fatto, nessuna - di queste preferenze sono condivise da entrambi i giocatori. Eppure questa versione di (1) non sembra significativamente diversa dalle altre. La situazione presenta ancora il carattere peculiare che ho messo in evidenza per mezzo dei miei undici esempi. Permetteremo quindi che nei problemi di coordinazione si presentino equilibri di non coordinazione.

	C1	C2	C3	C4
R1	si incontrano 1	0	0	0,5
R2	0	si incontrano 1	0	0,5
R3	0	0	si incontrano 1	0,5
R4	0,5	0,5	0,5	si incontrano 0,2

Figura 8

Tutti i miei problemi di coordinazione campione hanno due o più equilibri di coordinazione diversi. Questa molteplicità è importante in relazione al carattere peculiare dei problemi di coordinazione e andrebbe inclusa nella loro definizione. Se non c'è alcun considerevole conflitto d'interesse, il

problema di conseguire un unico equilibrio di coordinazione è più o meno banale. Esso sarà conseguito se la natura della situazione è chiara abbastanza perché ciascuno, date le sue aspettative, faccia la scelta migliore; perché ciascuno si attenda che ciascun altro, date le sue aspettative, faccia la scelta migliore, e così via. Come vedremo, queste condizioni non garantiscono la coordinazione se ci sono molteplici equilibri di coordinazione.

Molte delle situazioni con equilibrio di coordinazione unico sono ancora più banali (e meritano ancora di più di essere escluse). Per esempio, qualsiasi situazione in cui tutti gli agenti hanno davanti delle scelte (*strettamente*) *dominanti* - delle azioni che essi preferiscono indipendentemente dalle azioni degli altri - può avere un solo equilibrio (e, *a fortiori*, un solo equilibrio di coordinazione): la combinazione delle scelte dominanti. Una combinazione di scelte dominanti deve essere un equilibrio; ma potrebbe non essere un equilibrio di coordinazione, come nel noto Dilemma del Prigioniero, illustrato nella Figura 9, in cui R1 e C1 (la falsa confessione, nella storiella solita) sono dominanti e la loro combinazione  $\langle R1, C1 \rangle$  è un equilibrio di non coordinazione.

	C1	C2
R1	-8	-10
R2	-1	-2

Figura 9

Potremmo fare la congettura che ci sia dominanza in qualunque gioco di coordinazione pura con un unico equilibrio: che tutti, o almeno alcuni agenti abbiano davanti delle scelte dominanti, o almeno dominate. (Una scelta [*strettamente*] *dominata* è una scelta tale che, indipendentemente da quello che scelgono gli altri, voi avreste potuto fare una qualche altra scelta che si sarebbe rivelata migliore. Se una scelta è dominante, un'altra deve essere dominata; ma non viceversa, perché *quale* altra scelta sarebbe stata migliore per voi può dipendere dal modo in cui hanno scelto gli altri.) La congettura contiene questa parte di verità: in ogni gioco finito di coordinazione pura a due persone con un unico equilibrio almeno una azione di uno degli agenti è dominata. Dimostrazione:

Sia  $P(\langle R_j, C_k \rangle)$  la vincita nella combinazione  $\langle R_j, C_k \rangle$ , uguale per entrambi i giocatori.

Consideriamo un opportuno gioco con  $m$  righe e  $n$  colonne. Assumiamo, senza alcuna perdita di generalità, che le righe e le colonne siano disposte in modo tale che, per qualunque combinazione  $\langle R_i, C_1 \rangle$  sulla diagonale e per qualunque combinazione  $\langle R_j, C_k \rangle$  tale che  $j \geq i$  e  $k \geq i$ ,  $P(\langle R_j, C_k \rangle) \leq P(\langle R_i, C_i \rangle)$ . In particolare,  $\langle R_1, C_1 \rangle$  deve essere l'unico equilibrio, e  $P(\langle R_1, C_1 \rangle)$  deve essere superiore a ogni altra vincita del gioco.

Se  $\langle R_1, C_1 \rangle$  è la sola combinazione diagonale che è un massimo di riga o di colonna, allora  $R_m$  (se  $m \geq n$ ) o  $C_n$  (se  $n \geq m$ ) deve essere dominata.

Altrimenti, sia  $\langle R_a, C_a \rangle$ ,  $a \neq 1$ , la combinazione diagonale più a destra che è un massimo di riga o di colonna. Non è un massimo di entrambi i tipi poiché non si tratta di un equilibrio. Supponiamo, senza alcuna perdita di generalità, che sia un massimo di riga.

A meno che  $R_a$  sia strettamente dominata, c'è su  $R_a$  un massimo di colonna; sia  $\langle R_a, C_b \rangle$  quello più a destra.  $\langle R_a, C_b \rangle$  non è un massimo di riga perché non è un equilibrio; quindi  $P(\langle R_a, C_a \rangle) > P(\langle R_a, C_b \rangle)$ .

A meno che  $C_b$  sia strettamente dominata, su  $C_b$  c'è un massimo di riga; sia  $\langle R_{a'}, C_b \rangle$  quello più in basso. Poiché  $\langle R_a, C_b \rangle$  è un massimo di colonna,  $P(\langle R_a, C_b \rangle) \geq P(\langle R_{a'}, C_b \rangle)$ , e quindi  $P(\langle R_a, C_a \rangle) > P(\langle R_{a'}, C_b \rangle)$ .

A meno che  $R_{a'}$  sia strettamente dominata, su  $R_{a'}$  c'è un massimo di colonna; sia  $\langle R_{a'}, C_{b'} \rangle$  quello più a destra;  $P(\langle R_a, C_a \rangle) > P(\langle R_{a'}, C_{b'} \rangle)$ .

A meno che  $C_{b'}$  sia strettamente dominata, su  $C_{b'}$  c'è un massimo di riga; sia  $\langle R_{a''}, C_{b'} \rangle$  quello più in basso;  $P(\langle R_a, C_a \rangle) > P(\langle R_{a''}, C_{b'} \rangle)$ .

A meno che  $R_{a''}$  sia strettamente dominata, su  $R_{a''}$  c'è un massimo di colonna; sia  $\langle R_{a''}, C_{b''} \rangle$  quello più a destra;  $P(\langle R_a, C_a \rangle) > P(\langle R_{a''}, C_{b''} \rangle)$ . E così via.

Se  $\langle R_j, C_i \rangle$  è un massimo di colonna e  $P(\langle R_a, C_a \rangle) > P(\langle R_j, C_i \rangle)$ , allora  $\langle R_j, C_i \rangle$  è sopra la diagonale. Altrimenti infatti  $j \geq i$ , e quindi  $P(\langle R_j, C_i \rangle) \leq P(\langle R_i, C_i \rangle)$ . E siccome  $\langle R_j, C_i \rangle$  è un massimo di colonna,  $P(\langle R_j, C_i \rangle) = P(\langle R_i, C_i \rangle)$ . Ma allora anche  $\langle R_i, C_i \rangle$  è un massimo di colonna, e è a destra

di  $\langle Ra, Ca \rangle$  poiché  $P(\langle Ra, Ca \rangle) > P(\langle Ri, Ci \rangle)$ . Ma ciò è contrario alla nostra scelta di  $\langle Ra, Ca \rangle$ .

In particolare:  $\langle Ra, Cb \rangle$ ,  $\langle Ra', Cb' \rangle$ ,  $\langle Ra'', Cb'' \rangle$  ecc. sono sopra la diagonale.

Grazie a un'argomentazione analoga, se  $\langle Rj, Ci \rangle$  è un massimo di riga e  $P(\langle Ra, Ca \rangle) > P(\langle Rj, Ci \rangle)$ , allora  $\langle Rj, Ci \rangle$  è sotto la diagonale. In particolare:  $\langle Ra', Cb \rangle$ ,  $\langle Ra'', Cb' \rangle$  ecc. sono sotto la diagonale.

La sequenza di combinazioni che stavamo costruendo serpeggia quindi avanti e indietro attraverso la diagonale, come si vede nella Figura 10, cosicché  $a < a' < a'' \dots$  e  $b < b' < b'' \dots$ . Siccome il gioco è finito, queste sequenze hanno un termine, il che può accadere solo se una delle scelte  $Ra, Cb, Ra', Cb', Ra'', Cb''$  ecc. è strettamente dominata.

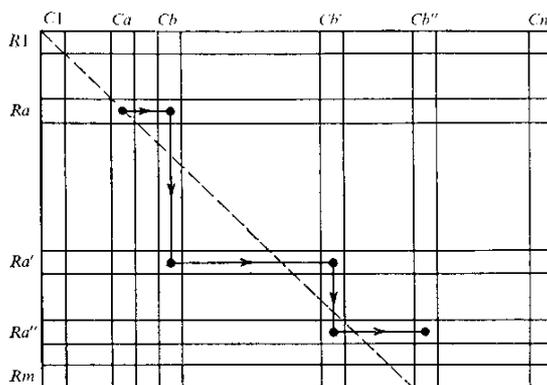


Figura 10

In un gioco finito di coordinazione pura a due persone con un unico equilibrio, la cancellazione di un'azione dominata produce un nuovo gioco, che è anch'esso un gioco finito di coordinazione pura a due persone con un unico equilibrio. La cancellazione può quindi essere ripetuta. Con successive cancellazioni di azioni dominate il gioco si trasforma in una situazione che è evidentemente banale perché i due giocatori hanno la possibilità di compiere una sola azione ciascuno. L'esito è determinato dal fatto che ciascuno ignora le azioni dominate, ciascuno si aspetta che ciascuno ignori le azioni dominate, e così via.

Il risultato che abbiamo appena dimostrato non può sfortunatamente venir rafforzato in nessuno dei modi che potremmo sperare. Non è estendibile

	C1	C2	C3	C4	C5	...
R1	64	0	0	0	0	0
R2	32	16	0	0	0	0
R3	0	8	4	0	0	0
R4	0	0	2	1	0	0
R5	0	0	0	0,5	0,25	
...	0	0	0	0,5	0,25	

Figura 11

L1	C1	C2		
R1	5	0	L2	C1
R2	0	1		
			R1	C2
			4	3
			1	2
			1	2

Figura 12

a giochi infiniti a due persone: la Figura 11 è un controesempio. Non è estendibile a giochi a tre persone: la Figura 12 è un controesempio. (Chiamiamo le scelte del terzo agente *livelli* L1 e L2; scriviamo le sue vincite al centro dei quadrati.) Non può venir rafforzato nel caso finito a due persone; la Figura 13 è un esempio con nessuna azione dominante e una sola azione dominata (e quell'unica è dominata soltanto da tutte le alternative insieme). Perciò non possiamo dire che la dominanza sia responsabile di tutti i casi di equilibri unici nei giochi di coordinazione pura.

Per escludere i casi banali, un problema di coordinazione deve avere più di un equilibrio di coordinazione. Ma questo requisito non è abbastanza forte. La figura 14 mostra due matrici in cui certamente ci sono molteplici equilibri di coordinazione (due a sinistra, quattro a destra). Eppure nessuno dei due agenti ha ancora bisogno di basare la sua scelta sulle sue aspettative

	C1	C2	C3
R1	16	0	0
R2	8	4	0
R3	0	2	1

Figura 13

	C1	C2
R1	1	1
R2	0	0

	C1	C2	C3
R1	1	1	0,2
R2	1	1	0,5
R3	0	0	0

Figura 14

circa la scelta dell'altro. Essi non hanno bisogno di cercare di raggiungere lo stesso equilibrio (di cercare di raggiungere la coordinazione) perché, se anche tentano di realizzare equilibri diversi, qualche equilibrio sarà comunque realizzato. Questi esempi presentano un altro genere di banalità, simile alla banalità dei casi con un unico equilibrio di coordinazione.

Una combinazione è un equilibrio se essa piace a ciascun agente *almeno quanto* qualunque altra combinazione che questo agente, date le scelte degli altri, avrebbe potuto conseguire. Chiamiamo una combinazione equilibrio *proprio* se essa piace a ciascun agente *più di* qualunque altra combinazione che, date le scelte degli altri, questo agente avrebbe potuto conseguire. In una matrice a due persone, per esempio, un equilibrio proprio è preferito dal giocatore di riga a tutte le altre combinazioni della sua colonna, e dal giocatore di colonna a tutte le altre combinazioni della sua riga. Nelle matrici della Figura 14 ci sono molteplici equilibri di coordinazione, ma sono tutti impropri.

	C1	C2	C3
R1	2	0	0
R2	0	2	0
R3	0	1	1

Figura 15

	C1	C2	C3
R1	1	0	0
R2	0	1	1
R3	0	1	1

	C1	C2	C3	C4
R1	1	1	0	0
R2	1	1	0	0
R3	0	0	1	1
R4	0	0	1	1

Figura 16

Non è necessario stipulare che tutti gli equilibri di un problema di coordinazione debbano essere propri; sembra che la matrice della Figura 15 debba essere considerata essenzialmente simile ai nostri chiari esempi di problemi di coordinazione, nonostante l'atipicità del suo equilibrio  $\langle R3, C3 \rangle$ . I due equilibri di coordinazione propri  $\langle R1, C1 \rangle$  e  $\langle R2, C2 \rangle$  sono sufficienti a rendere non banale il problema. Stipulerò invece che un problema di coordinazione debba contenere almeno due equilibri di coordinazione propri.

Questa è soltanto una - la più forte - delle molte restrizioni giustificabili. Potremmo preferire una restrizione più debole che non escludesse matrici come quelle della Figura 16. Ma una restrizione soddisfacente sarebbe complicata e comporterebbe troppe restrizioni ulteriori. E situazioni come quella della Figura 16 possono egualmente essere recuperate con la restrizio-

	C1	C2'
R1	1	0
R2'	0	1

Figura 17

	C1	C2	C3
R1	1	1	0
R2	0	1	1
R3	0	0	1

Figura 18

ne forte che abbiamo adottato. Sia R2' la disgiunzione di R2 e R3 e C2' la disgiunzione di C2 e C3 nella matrice di sinistra. La stessa situazione può allora essere rappresentata dalla nuova matrice della Figura 17, che ha due equilibri di coordinazione propri. La matrice di destra può essere contratta in modo analogo. Ma le matrici del tipo di quella della Figura 18, che sono escluse dalle restrizioni forti, e effettivamente andrebbero escluse, non possono venir recuperate per mezzo di contrazioni del genere.

Per riassumere: i problemi di coordinazione - situazioni che assomigliano ai miei undici esempi per quegli importanti aspetti per i quali essi si assomigliano reciprocamente<sup>4</sup> - sono situazioni di decisione interdependente da parte di due o più agenti in cui predomina la coincidenza di interessi e in cui ci sono due o più equilibri di coordinazione propri. Si potrebbe anche dire benché in modo meno informativo di quanto si potrebbe pensare che sono situazioni in cui, relativamente a *qualche* classificazione delle azioni,

<sup>4</sup>Cfr. Michael Slote, "The Theory of Important Criteria", *Journal of Philosophy*, 63 (1966), pp. 211-224. Slote mostra come noi introduciamo comunemente una classe per mezzo di esempi e riteniamo che le caratteristiche definitorie della classe siano quei tratti distintivi dei nostri esempi che sembrano importanti ai fini di una comprensione della loro natura. È questo che presumo di fare io stesso qui e altrove.

gli agenti hanno l'interesse comune di compiere tutti la stessa azione tra più azioni alternative.

### 1.3 Risoluzione dei problemi di coordinazione

Gli agenti che si trovano di fronte a un problema di coordinazione possono riuscire o non riuscire a comportarsi ciascuno in modo tale da raggiungere uno dei possibili equilibri di coordinazione. Può darsi che vi riescano per pura fortuna, sebbene alcuni di essi scelgano senza tener conto delle azioni che si aspettano dagli altri (comportandosi così forse perché non riescono a indovinare che cosa faranno gli altri, o forse perché le probabilità di coordinazione sembrano tanto piccole da essere trascurabili). Ma è più probabile che vi riescano - se riescono - grazie all'azione di un sistema di reciproche aspettative opportunamente concordanti. Così, nell'esempio (1) può darsi che io vada in un certo posto perché mi aspetto che tu vada là, mentre tu vai là perché ti aspetti che ci vada io; nell'esempio (2) può darsi che io richiami perché mi aspetto che tu non lo faccia, mentre tu non richiami perché ti aspetti che lo faccia io; nell'esempio (4) può darsi che ciascuno di noi guidi a destra perché si aspetta che gli altri facciano altrettanto; e così via. In generale, può darsi che ciascuno svolga la sua parte in uno dei possibili equilibri di coordinazione perché si aspetta che gli altri svolgano la loro, raggiungendo così quell'equilibrio.

Se un agente avesse una fiducia completa nella sua aspettativa che gli altri svolgano la loro parte in un certo equilibrio di coordinazione proprio, egli avrebbe una ragione decisiva per svolgere la propria parte. Ma se, come in qualunque caso reale, la sua fiducia è meno che completa, egli deve confrontare la preferenza a svolgere la propria parte nel caso che gli altri svolgano la loro con le preferenze a comportarsi diversamente nel caso che gli altri non la svolgano. Egli ha una ragione decisiva per svolgere la propria parte se ha una fiducia *sufficiente* nell'aspettativa che gli altri svolgano la loro. Il grado di fiducia che è sufficiente dipende da tutte le sue vincite e a volte dalle probabilità relative che egli assegna ai diversi *modi* in cui gli altri potrebbero non svolgere tutti la propria parte, nel caso che non tutti la svolgano. Per esempio, nel problema di coordinazione illustrato nella Figura 19, il giocatore di riga dovrà svolgere la propria parte nell'equilibrio di coordinazione  $\langle R1, C1 \rangle$ , scegliendo R1, se la sua fiducia che il giocatore di colonna svolgerà la sua parte, scegliendo C1, sarà superiore a 0,5. Ma nei problemi di coordinazione illustrati nella Figura 20 il giocatore di riga dovrà scegliere R1 solo se la sua fiducia che il giocatore di colonna scelga

	C1	C2
R1	1	0
R2	0	1

Figura 19

	C1	C2
R1	1	-8
R2	0	1

	C1	C2
R1	1	0
R2	0	9

	C1	C2
R1	3	-26
R2	0	1

Figura 20

C1 è superiore a 0,9. Se, per esempio, la sua fiducia che il giocatore di colonna scelga C1 è di 0,8, egli farà meglio a scegliere R2, sacrificando la sua possibilità di conseguire la coordinazione in  $\langle R1, C1 \rangle$  al fine di far fronte alla possibilità che la sua aspettativa fosse sbagliata. E nel problema di coordinazione illustrato nella Figura 21 può darsi che il giocatore di riga sia sicuro che, se il giocatore di colonna non svolge la sua parte di  $\langle R1, C1 \rangle$ , perlomeno sceglierà C2, e non C3; se è così, il giocatore di riga dovrà scegliere R1 se la sua fiducia che il giocatore di colonna sceglierà C1 è superiore a 0,5. Oppure può darsi che il giocatore di riga pensi che, se il giocatore di colonna non sceglie C1, è altrettanto probabile che scelga C3 quanto C2; se è così, il giocatore di riga dovrà scegliere R1 solo se la sua fiducia che il giocatore di colonna sceglierà C1 è superiore a 0,9. Oppure può darsi che il giocatore di riga sia sicuro che, se il giocatore di colonna non sceglie C1, sceglierà invece C3; se è così, il grado di fiducia minimo

	C1	C2	C3
R1	1	0	-8
R2	0	1	9

Figura 21

sufficiente del giocatore di riga è di circa 0,95. L'intensità dell'aspettativa concordante necessaria per produrre la coordinazione in un certo equilibrio è una misura della difficoltà di conseguire la coordinazione, poiché, comunque vengano prodotte le aspettative concordanti, le aspettative più deboli saranno prodotte più facilmente di quelle più forti. (Possiamo immaginare casi in cui, per conseguire la coordinazione in un certo equilibrio, è richiesto un tale grado di fiducia reciproca che il successo è impossibile. Immaginiamo che un milionario faccia l'offerta di distribuire equamente la propria fortuna tra mille persone se ciascuna gli manda dieci dollari; se anche una sola persona non glieli manderà, il milionario si terrà tutto quello che gli è stato mandato. Presumo che, indipendentemente da quello che i mille fanno per accrescere la loro fiducia reciproca, sia praticamente certo che il milionario non dovrà pagare. Così, se sono uno dei mille, mi terrò i miei dieci dollari.)

Possiamo raggiungere la coordinazione agendo sulle nostre aspettative concordanti che concernono le nostre azioni reciproche. E possiamo acquisire quelle aspettative, o correggere o rafforzare qualunque aspettativa che abbiamo già, facendo del nostro meglio per metterci nei panni degli altri. Se io so qual è il vostro parere sui dati di fatto che determinano gli effetti probabili delle vostre azioni alternative, e se so quali preferite tra i risultati possibili e so che possedete un minimo di razionalità pratica, allora sono in grado di ricostruire il vostro ragionamento pratico per calcolare che cosa probabilmente farete, in modo da potermi comportare adeguatamente.

Nel caso di un problema di coordinazione, o di qualsiasi altro problema di decisione interdipendente, uno dei dati di fatto che determinano gli effetti probabili delle vostre azioni alternative è la mia azione. Al fine di calcolare che cosa voi farete ricostruendo il vostro ragionamento pratico, io ho bisogno di calcolare che cosa *voi* vi aspettate che *io* faccia.

Io so che, come io sto cercando di calcolare che cosa farete voi ricostruendo il vostro ragionamento pratico, così può darsi che voi stiate cercando di calcolare che cosa farò io ricostruendo il mio ragionamento. Anche questo, come qualunque altra cosa che voi potreste fare per calcolare che cosa io farò, fa parte del vostro ragionamento. Così, per ricostruire il vostro ragionamento, può darsi che io debba ricostruire il vostro tentativo di ricostruire il mio ragionamento.

Ma non è finita qui. Io posso ragionevolmente aspettarmi che voi vi rendiate conto che può darsi che io, a meno che sappia già quello che vi aspettate che faccia, debba cercare di ricostruire il vostro tentativo di ricostruire il mio ragionamento. Così io posso aspettarmi che voi cerchiate di ricostruire il mio tentativo di ricostruire il vostro tentativo di ricostruire il mio ragionamento. Può dunque darsi che il mio ragionamento debba includere il tentativo di ricostruire il vostro tentativo di ricostruire il mio tentativo di ricostruire il vostro tentativo di ricostruire il mio ragionamento. E così via. Prima che la questione ci sfugga di mano sarà utile introdurre il concetto di *aspettativa di ordine superiore*, definito ricorsivamente nel modo seguente:

Un'aspettativa di primo ordine circa qualcosa è una aspettativa ordinaria circa quella cosa.

Un'aspettativa di ordine  $(n + 1)$  circa qualcosa  $(n \geq 1)$  è un'aspettativa ordinaria circa l'aspettativa di ordine  $n$  di qualcun altro circa quella cosa.

Per esempio, se io mi aspetto che tu ti aspetti che tuoni, ho un'aspettativa di secondo ordine che tuoni.

Ogniqualevolta ricostruisco un vostro ragionamento pratico, le mie aspettative di secondo ordine circa i dati di fatto, unite alle mie aspettative di primo ordine circa le vostre preferenze e la vostra razionalità, giustificano la mia elaborazione di aspettative di primo ordine circa la vostra azione. Nel caso di problemi di decisione interdipendente - per esempio, problemi di coordinazione - alcune delle aspettative di secondo ordine richieste devono riguardare la mia stessa azione.

Consideriamo il nostro primo esempio di problema di coordinazione: una situazione in cui tu e io vogliamo incontrarci andando nello stesso posto. Supponiamo che, dopo averci riflettuto, io decida di recarmi in un certo luogo. Il ragionamento pratico fondamentale che mi porta a questa scelta è illustrato nella Figura 22. (In tutti i diagrammi di questo tipo le frecce in grassetto rappresentano implicazioni, mentre le altre rappresentano connessioni causali tra gli stati mentali o tra le azioni di un agente razionale.)

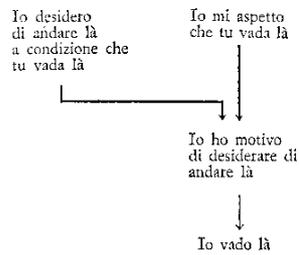


Figura 22

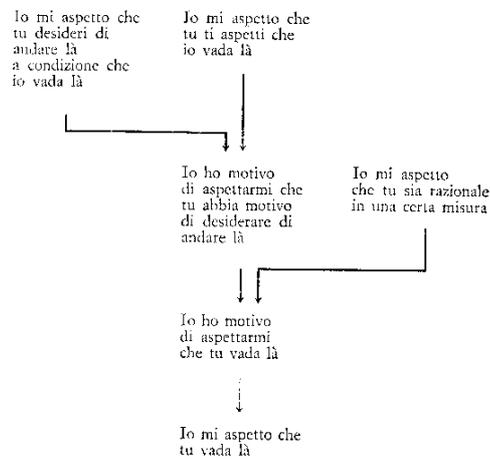


Figura 23

Se la mia premessa a questo ragionamento, cioè la mia aspettativa che tu vada là, era ottenuta ricostruendo il tuo ragionamento, la mia ricostruzione è illustrata nella Figura 23. Se la mia premessa a questa ricostruzione, cioè la mia aspettativa che tu ti attenda che io vada là, era ottenuta ricostruendo la tua ricostruzione del mio ragionamento, la mia ricostruzione della tua ricostruzione è illustrata nella Figura 24. E così via. Il mio ragionamento complessivo (semplificato tralasciando le premesse di razionalità) può essere rappresentato come nella Figura 25, quale che sia il numero finito di stadi che può essere richiesto dal mio uso di qualsivoglia aspettativa di ordine superiore, circa le nostre azioni e le nostre preferenze condizionali, che possa essermi disponibile. Le ricostruzioni sono incastrate l'una nell'altra fino a una certa profondità finita: il mio ragionamento (linea tratteggiata più esterna) contiene una ricostruzione del tuo (linea tratteggiata succes-

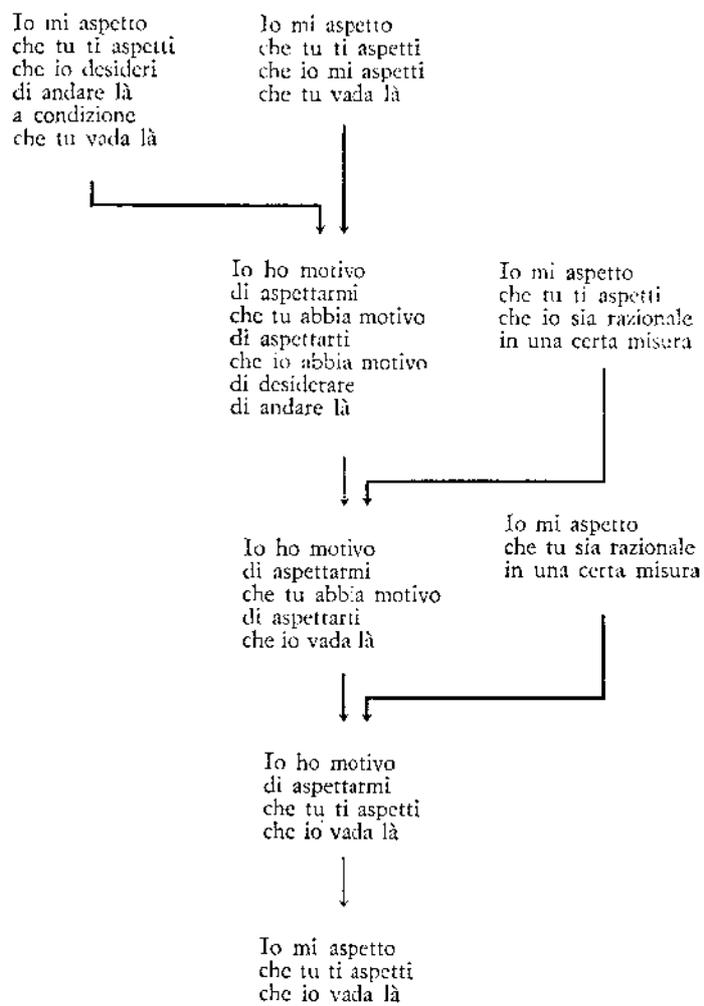


Figura 24

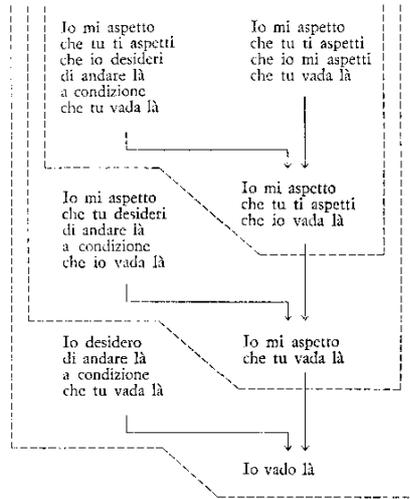


Figura 25

siva), che contiene una ricostruzione della tua ricostruzione del mio (linea tratteggiata successiva), e così via.

Così, se mi trovo ad avere una aspettativa di n-esimo ordine circa l'azione in questo problema di coordinazione a due persone, posso procedere, attraverso le ricostruzioni incastrate l'una nell'altra, verso l'esterno, fino a raggiungere aspettative di ordine sempre più basso concernenti l'azione. Se vado avanti abbastanza, e se sono disponibili tutte le necessarie aspettative di ordine superiore circa le preferenze e la razionalità, alla fine arrivo a un'aspettativa di primo ordine circa la tua azione - che è ciò che mi occorre per sapere come devo comportarmi.

È chiaro che un processo di ricostruzione simile è possibile nei problemi di coordinazione con più di due agenti. In generale, le mie aspettative di ordine superiore circa qualcosa sono le mie aspettative circa le aspettative di  $x_1$  circa le aspettative di  $x_2...$  circa quella cosa. (La sequenza  $x_1, x_2...$  può ripetersi, ma io non posso essere  $x_1$  e nessuno può occorrere due volte in successione immediata.) Così, quando ci sono  $m$  agenti io posso avere esattamente  $(m-1)^n$  aspettative di n-esimo ordine diverse, corrispondenti alle  $(m-1)^n$  diverse sequenze ammissibili di lunghezza  $n$ . La ricostruzione in generale è ramificata: è eseguita a partire dagli stadi in cui  $m-1$  fra le mie varie aspettative di ordine  $(n+1)$ -esimo circa l'azione, più le premesse ausiliarie, danno una delle mie aspettative di ordine n-esimo circa l'azione. Io ho soppresso la ramificazione ponendo  $m=2$ , ma il caso generale è in linea

di principio lo stesso.

Si noti che la ricostruzione *non* è un'interazione tra persone. È un processo in cui una persona calcola le conseguenze delle sue credenze relative al mondo - un mondo che egli crede contenga altre persone che calcolano le conseguenze delle proprie credenze, tra le quali c'è la loro credenza in altre persone che... Grazie alla nostra interazione nel mondo noi acquisiamo diverse aspettative di ordine superiore che possono servirci come premesse. Nel nostro successivo ragionare siamo monadi prive di finestre che fanno del loro meglio per rispecchiarsi reciprocamente, per rispecchiare reciprocamente il loro reciproco specchiarsi, e così via.

Naturalmente io non penso che uno risolva un problema di coordinazione prima prendendo da qualche parte un'aspettativa di diciassettesimo ordine e poi sedendosi a fare le sue ricostruzioni. Da un lato, noi abbiamo raramente delle aspettative di ordine superiore, diciamo, al quarto. D'altro lato, qualunque situazione ordinaria che possa giustificare un'aspettativa di ordine superiore giustificherebbe anche direttamente aspettative di ordine inferiore, senza ricorso a ricostruzioni incastrate l'una nell'altra.

Nondimeno, date le premesse ausiliarie occorrenti, un'aspettativa di ordine arbitrariamente elevato concernente l'azione fornisce effettivamente all'agente *una* buona ragione per scegliere un'azione. Può darsi, e normalmente sarà così, che si tratti di una ragione tra le molte che, insieme, sono sufficienti a giustificare la sua scelta. Supponiamo che l'agente abbia un qualche motivo iniziale per avere aspettative di vari ordini circa le azioni proprie e dei suoi partner. E supponiamo che le premesse ausiliarie siano disponibili. Ciascuna delle sue aspettative originarie gli fornisce indipendentemente un motivo per comportarsi in un modo o nell'altro. Se è fortunato, tutti questi motivi indipendenti saranno motivi per compiere la stessa azione.<sup>5</sup> In

---

<sup>5</sup>Michael Scriven, in "An Essential Unpredictability in Human Behavior", *Scientific Psychology: Principles and Approaches*, a cura di B.B. Wolman, Basic Books, New York, 1936, ha discusso la ricostruzione reciproca del ragionamento pratico tra agenti che partecipano a un gioco di conflitto e che non vogliono conformarsi alle loro aspettative reciproche. C'è un avvicendamento ciclico: dalla mia aspettativa di  $(n + 4)$ -esimo ordine di andare a Minsk alla mia aspettativa di  $(n + 3)$ -esimo ordine che tu vada a Pinsk, alla mia aspettativa di  $(n + 2)$ -esimo ordine di andare a Pinsk, alla mia aspettativa di  $(n + 1)$ -esimo ordine che tu vada a Minsk, alla mia aspettativa di  $n$ -esimo ordine di andare a Minsk... Scriven osserva che non possiamo agire entrambi sulla base di ricostruzioni complete e accurate del nostro ragionamento reciproco. Egli ritiene che questo dimostri l'imprevedibilità umana. Ma forse dimostra semplicemente che gli agenti non possono disporre entrambi di un tempo sufficiente a portare a termine le loro ricostruzioni, in quanto il tempo di cui l'uno ha bisogno aumenta in relazione al tempo che l'altro impiega. Cfr. David Lewis e Jane Richardson, "Scriven on Human Unpredictability", *Philosophical Studies*, 17 (1966) pp. 69-74.

questo caso l'azione è fortemente giustificata, in quanto giustificata in modo ridondante; egli ha più motivi di compierla di quanti avrebbero potuto essere forniti isolatamente da una qualunque delle sue aspettative originarie.

Prima ho detto che la coordinazione può essere raggiunta razionalmente con l'aiuto di reciproche aspettative concordanti circa l'azione. Abbiamo visto che queste possono essere derivate da aspettative di primo ordine e di ordine superiore circa l'azione, le preferenze e la razionalità. Così generalizziamo: la coordinazione può essere conseguita razionalmente con l'aiuto di un sistema di aspettative reciproche concordanti, di primo ordine e di ordini superiori, circa le azioni, le preferenze e la razionalità degli agenti.

Più gli ordini dell'aspettativa circa l'azione contribuiscono a una decisione dell'agente, più giustificazioni indipendenti avrà l'agente; e quanto più egli è cosciente di quelle giustificazioni, tanto più saldamente sarà determinata la sua scelta. Le circostanze favorevoli alla risoluzione di un problema di coordinazione sono quindi circostanze in cui gli agenti vengono motivati a elaborare aspettative reciproche appartenenti a un sistema armonioso. E più sono gli ordini meglio è.

Nel considerare i modi di risolvere i problemi di coordinazione ho rimandato la risposta che viene in mente per prima: per accordo. Se gli agenti possono comunicare (a un costo non eccessivo), essi possono assicurarsi una visione comune del loro problema discutendolo. Possono scegliere un equilibrio di coordinazione: uno arbitrario, oppure uno particolarmente vantaggioso per qualcuno di essi o per tutti, oppure uno che essi possono conseguire senza troppa fiducia reciproca. E ciascuno può assicurare gli altri che svolgerà la propria parte nell'equilibrio prescelto. La coordinazione per mezzo di accordo non è, naturalmente, alternativa alla coordinazione per mezzo di aspettative reciproche concordanti. Caso mai, l'accordo è un mezzo per produrre quelle aspettative. È un mezzo particolarmente efficace, in quanto produce forti aspettative concordanti di svariati ordini.

Supponiamo che tu e io vogliamo incontrarci domani; oggi ci incontriamo per caso e ci diamo un appuntamento. Con questo ciascuno manifesta il suo interesse ad andare dove va l'altro e la sua intenzione di andare in un certo luogo. Considerando tale manifestazione noi elaboriamo delle aspettative concordanti di primo ordine circa le preferenze e l'azione reciproche. Considerando la reciproca considerazione di essa, possiamo anche elaborare aspettative concordanti di secondo ordine. Considerando la reciproca considerazione della reciproca considerazione di essa, possiamo addirittura elaborare aspettative concordanti di terzo ordine. E così via; non all'infinito, naturalmente, ma limitatamente alla quantità di ragionamenti che facciamo e alla quantità che ci attribuiamo reciprocamente - probabilmente

per uno o due passaggi ancora. Il risultato è un sistema di aspettative reciproche concordanti di diversi ordini, tendenti alla coordinazione per mezzo di ricostruzione.

L'accordo degli agenti potrebbe essere uno scambio di promesse formali o tacite. Ma non è necessario che lo sia. Anche un uomo fedele alla parola data può sciogliere il vincolo della promessa con una ritrattazione esplicita, se non in altro modo. Uno scambio di dichiarazioni delle intenzioni attuali sarà sufficiente, anche se ciascuno ritiene esplicitamente suo diritto cambiare in seguito i propri progetti. Nessuno ha bisogno di impegnarsi ad agire contro il proprio interesse. Semmai, sarà nell'interesse di ciascuno fare esattamente ciò che egli ha indotto gli altri ad aspettarsi da lui, poiché quell'azione sarà per lui la migliore se gli altri agiscono sulla base delle loro aspettative.

Se uno si considera vincolato da una promessa ha un secondo incentivo indipendente. Le sue vincite risultano modificate, in quanto a tutte le scelte tranne una egli ha associato la responsabilità dell'infrazione di una promessa. Anzi, può darsi che, con la promessa, egli modifichi le sue vincite al punto che la situazione non sia più un problema di coordinazione. Per esempio, l'azione promessa dell'agente potrebbe diventare la sua scelta dominante: egli potrebbe voler mantenere la promessa comunque, che ne risulti o no la coordinazione. Se una tale promessa solenne è fatta pubblicamente, gli altri sapranno che devono conformarsi a quello che costui ha promesso, perché sanno quello che farà. Tali promesse rigide sono un modo di sbarazzarsi dei problemi di coordinazione, non un modo di risolverli.

L'accordo esplicito è un mezzo particolarmente efficace e diffuso per ottenere la coordinazione - tanto che siamo tentati di chiamare accordo *tacito* la coordinazione ottenuta diversamente. Ma l'accordo (nel senso letterale del termine) non è l'unica fonte di aspettative concordanti che ci possa aiutare a risolvere i problemi di coordinazione. Facciamo liberamente a meno dell'accordo se ci troviamo già soddisfatti del contenuto e dell'intensità delle nostre aspettative reciproche. Facciamo necessariamente a meno di esso se non abbiamo alcuna possibilità di comunicare, o se possiamo comunicare solo a un prezzo che supera le nostre aumentate probabilità di coordinazione (per esempio, se siamo dei cospiratori pedinati).

Schelling ha fatto esperimenti con problemi di coordinazione in cui gli agenti non possono comunicare. I suoi soggetti sanno soltanto di avere una nozione comune del problema: per esempio, possono ricevere istruzioni che descrivono il problema e che dicono che ciascuno riceve le stesse istruzioni. Risulta che, in una situazione sperimentale, soggetti particolarmente educati possono spesso riuscire molto bene molto meglio di quanto sarebbe probabile a risolvere, senza comunicare, problemi di coordinazione nuovi.

Essi ricercano un equilibrio di coordinazione che sia in qualche modo *saliente*: un equilibrio che si distingua dagli altri in quanto unico per qualche aspetto rilevante. Non è necessario che esso sia *buono* in modo unico; anzi, potrebbe essere cattivo in modo unico. Esso deve semplicemente essere unico in un qualche modo che i soggetti noteranno, che essi si aspettano di notare reciprocamente, e così via. Se equilibri di coordinazione diversi sono unici secondo modalità rilevanti diverse, i soggetti dovranno assomigliarsi per quanto riguarda l'importanza relativa che attribuiscono ai diversi termini del confronto; ma spesso essi si assomigliano abbastanza per risolvere il problema.

Come possiamo spiegare la coordinazione per salienza? Potrebbe darsi che i soggetti tendano tutti a scegliere ciò che è saliente come estrema risorsa, quando non hanno alcun motivo di scelta più forte. Oppure potrebbe darsi che essi si aspettino reciprocamente di avere quella tendenza e si comportino di conseguenza; oppure potrebbe darsi che essi si aspettino reciprocamente di aspettarsi reciprocamente di avere quella tendenza e di comportarsi di conseguenza, e si comportino di conseguenza; e così via. Oppure più probabilmente potrebbe esserci una mescolanza di questi motivi. Le loro aspettative di primo ordine e di ordine superiore di una tendenza a scegliere, come estrema risorsa, ciò che è saliente formerebbero un sistema di aspettative concordanti capace di fissare la coordinazione sull'equilibrio saliente.

Se le loro aspettative hanno prodotto effettivamente la coordinazione, non importerà se qualcuno ha davvero scelto ciò che è saliente come estrema risorsa. Ciascuno infatti avrà avuto un buon motivo per fare la sua scelta, per cui la sua scelta non sarà stata un'estrema risorsa.

Così, anche in un problema di coordinazione nuovo, che è un caso estremo, gli agenti possono a volte ottenere senza comunicare le aspettative concordanti di cui hanno bisogno. Un caso più facile, e più comune, è quello di un problema di coordinazione *familiare* senza comunicazione. Qui la fonte delle aspettative reciproche degli agenti è il precedente: la conoscenza degli esempi, risolti in precedenza, del problema di coordinazione attuale.

## 1.4 Convenzione

Partiamo dal caso più semplice di coordinazione in base al precedente e generalizziamo in vari modi. Incontreremo così il fenomeno che io chiamo *convenzione*, argomento di questo libro.

Supponiamo che ci sia stato sottoposto un problema di coordinazione, e

di aver raggiunto un equilibrio di coordinazione abbastanza buono. Dato di nuovo esattamente lo stesso problema, forse ciascuno di noi ripeterà quello che ha fatto prima. Se ciò avviene, raggiungeremo la stessa soluzione. Se tu e io ci siamo incontrati ieri - per fortuna, per accordo, per salienza, o in qualunque altro modo - e oggi ci troviamo a doverci incontrare di nuovo, potremmo tornare entrambi nel luogo di incontro di ieri, sperando ciascuno di trovare là l'altro. Se la nostra conversazione telefonica è stata interrotta e tu hai richiamato mentre io aspettavo, allora, se siamo interrotti di nuovo durante la stessa telefonata, io aspetterò di nuovo.

Possiamo spiegare l'efficacia del precedente nello stesso modo in cui abbiamo spiegato quella della salienza. In realtà, il precedente è semplicemente la fonte di un importante tipo di salienza: la rilevante unicità di un equilibrio dovuta al fatto che l'abbiamo conseguito l'ultima volta. Se non abbiamo forti motivi per comportarci diversamente, può darsi che tendiamo a ripetere l'azione che ci è riuscita in precedenza. Sia che noi abbiamo realmente questa tendenza sia che non l'abbiamo, possiamo in qualche modo aspettarci reciprocamente di averla, o aspettarci reciprocamente di aspettarci reciprocamente di averla, e così via: in altri termini, possiamo avere ciascuno aspettative di primo ordine e di ordine superiore secondo le quali gli altri svolgeranno la loro parte nel vecchio equilibrio di coordinazione, a meno che abbiano motivo di comportarsi diversamente. L'aspettativa, da parte di ciascuno, che gli altri svolgano la loro parte, rafforzata eventualmente da una ricostruzione che utilizza le nostre aspettative di ordine superiore, fornisce a ciascuno qualche motivo per svolgere la propria parte. E, se le sue aspettative iniziali dell'uno o dell'altro ordine erano abbastanza forti, egli avrà un motivo decisivo per svolgere la propria parte. E così la svolgerà.

Ho supposto che ci sia presentato un problema di coordinazione, e che poi ci sia presentato di nuovo lo stesso problema. Ma, naturalmente, potrebbe darsi che un identico problema non ci venga mai sottoposto due volte. Ci deve essere per lo meno questa differenza: la seconda volta possiamo far ricorso alla nostra esperienza della prima volta. Più in generale, i due problemi saranno differenti sotto diversi aspetti indipendenti. Non possiamo fare esattamente quello che abbiamo fatto prima. Niente di ciò che potremmo fare questa volta è esattamente uguale a quello che abbiamo fatto prima (uguale ad esso sotto ogni aspetto), in quanto le situazioni non sono esattamente uguali.

Supponiamo, quindi, non che ci venga sottoposto di nuovo il problema iniziale, ma piuttosto che ci venga presentato un nuovo problema di coordinazione analogo in qualche modo a quello iniziale. Guidati da qualunque analogia notiamo, noi tendiamo a seguire il precedente cercando di conse-

guire, nel nuovo problema, un equilibrio di coordinazione che corrisponda in modo unico a quello che abbiamo conseguito in precedenza.

Potrebbero esserci analogie alternative. Se è così, c'è spazio per una certa ambiguità relativa a che cosa vuol dire seguire il precedente e fare quello che si è fatto in precedenza. Supponiamo che ieri io ti abbia chiamato al telefono e che ti abbia richiamato quando siamo stati interrotti. Oggi tu chiami me e veniamo interrotti. Abbiamo un precedente in cui io richiamavo e un precedente - lo stesso - in cui chi aveva chiamato richiamava. Ma questa volta chi ha chiamato sei tu. Indipendentemente da quello che faccio stavolta, faccio qualcosa di analogo a quello che ho fatto in precedenza. Il nostro precedente ambiguo non ci aiuta.

Di fatto, ci sono sempre innumerevoli analogie alternative. Se non fosse per il fatto che siamo uniformemente portati a notare alcune analogie e a ignorarne altre - quelle che chiamiamo, rispettivamente, "naturali" e "artificiose" -, i precedenti sarebbero sempre completamente ambigui e privi di valore. Ogni equilibrio di coordinazione del nostro nuovo problema (e anche ogni altra combinazione) corrisponde in modo unico a quello che abbiamo fatto prima in virtù di qualche analogia, ha in comune con esso soltanto *qualche* descrizione caratteristica. Fortunatamente, la maggior parte delle analogie sono artificiose. Noi le ignoriamo; non siamo inclini a permetter loro di guidare la nostra scelta, né ci aspettiamo reciprocamente di avere una tale tendenza, né ci aspettiamo reciprocamente di aspettarci reciprocamente di avere una tale tendenza, e così via. E fortunatamente abbiamo appreso che noi tutti noteremo per lo più le stesse analogie. per questo motivo che, in pratica, i precedenti possono essere, e spesso sono, non ambigui. Se noi notiamo soltanto una delle analogie che sussistono tra il precedente e il nostro problema, o se una delle analogie che notiamo ci sembra molto più rilevante delle altre, o se, anche, molte sono rilevanti ma concordano tutte nell'indicare la stessa scelta, allora le altre analogie non importano. Noi non siamo in difficoltà fino a che analogie discordanti non si propongono alla nostra attenzione.

Più sono gli aspetti sotto i quali il nuovo problema e il precedente sono simili, maggiore è la probabilità che analogie diverse si rivelino concordanti, minore sarà lo spazio per l'ambiguità, e più facile sarà conformarsi al precedente. Un precedente in cui io, che avevo chiamato per primo, richiamavo è ambiguo rispetto a un nuovo problema in cui chi ha chiamato per primo sei tu, ma non rispetto a un nuovo problema in cui chi ha chiamato per primo sono di nuovo io. È per questo che ho posto inizialmente la condizione che il nuovo problema fosse simile al precedente sotto tutti gli aspetti.

In generale, la salienza è unicità di un equilibrio di coordinazione sotto

un aspetto particolarmente rilevante. La salienza dovuta al precedente non fa eccezione: è unicità di un equilibrio di coordinazione in virtù della sua analogia particolarmente rilevante con ciò che è stato fatto con successo in precedenza.

Finora ho supposto che gli agenti che hanno istituito il precedente siano gli stessi che si conformano a esso. Questo ci garantiva che gli agenti ai quali veniva sottoposto il nuovo problema fossero a conoscenza delle particolarità e dell'esito del primo e si aspettassero reciprocamente di esserlo, si aspettavano reciprocamente di aspettarsi reciprocamente di esserlo, e così via. Ma questo non è un sistema infallibile, né è l'unico. Per esempio, se ieri ti ho raccontato di due che si sono separati nella metropolitana e poi si sono incontrati di nuovo in Charles Street, e oggi noi ci separiamo nello stesso modo, può darsi che decidiamo indipendentemente di andare a aspettarci in Charles Street. Non è affatto rilevante che quello che ti ho raccontato fosse vero, o che tu pensassi che lo fosse, o che io pensassi che lo fosse, o addirittura che io lo presentassi come tale. Un precedente fittizio sarebbe capace quanto uno reale di suggerirci una linea d'azione, e quindi sarebbe una fonte altrettanto valida di aspettative reciproche concordanti, capaci di farci incontrare. Perciò stabiliremo soltanto che gli agenti che hanno dinanzi il nuovo problema siano in qualche modo a conoscenza del precedente, si aspettino reciprocamente di esserne a conoscenza, e così via.

Finora ho supposto che il precedente da seguire sia singolo. Ma potrebbero essercene parecchi. Potremmo essere tutti a conoscenza di una classe di problemi di coordinazione precedenti, analoghi in modo naturale al nostro problema presente e l'uno all'altro, problemi in cui venivano raggiunti equilibri di coordinazione analoghi. Ciò equivale a dire che le azioni degli agenti si conformavano a una qualche notevole regolarità. Siccome il nostro attuale problema è opportunamente analogo ai precedenti, possiamo raggiungere un equilibrio di coordinazione conformandoci tutti a questa stessa regolarità. Ciascuno di noi desidera conformarsi a essa se lo fanno gli altri; ciascuno ha una *preferenza condizionale* per la conformità. Se ci conformiamo, la spiegazione è del solito tipo: se non c'è alcuna ragione particolare per comportarsi altrimenti, noi tendiamo a seguire il precedente; ci aspettiamo reciprocamente di avere questa tendenza; ci aspettiamo reciprocamente di aspettarcela; e così via. Abbiamo le nostre aspettative di primo ordine e di ordine superiore concordanti, ed esse ci danno la possibilità di raggiungere un equilibrio di coordinazione.

Non importa il *perché* la coordinazione sia stata conseguita, nei casi precedenti, in equilibri analoghi. Se anche si fosse realizzata per caso, potremmo ancora seguire l'insieme precedente. Una verosimile successione di

eventi potrebbe essere questa: il primo caso, o i primi casi, hanno funzionato come precedenti per il successivo, questi per il successivo, e così via. In modo analogo, indipendentemente da come si sono verificati i nostri precedenti, seguendoli questa volta noi aggiungiamo questo caso all'insieme dei precedenti d'ora in poi disponibili.

Parecchi precedenti sono meglio di uno, non solo perché noi apprendiamo per ripetizione, ma anche perché le differenze tra i precedenti ci aiutano a risolvere l'ambiguità. Se anche la nostra situazione attuale presenta analogie naturali discordanti con un qualunque precedente, probabilmente solo una di queste analogie sarà valida tra i precedenti; per cui noi presteremo attenzione a quella soltanto. Supponiamo di conoscere diversi casi in cui una conversazione telefonica interrotta veniva ripristinata, e supponiamo che, in tutti i casi, a richiamare fosse colui che aveva chiamato per primo. In alcuni casi chi aveva chiamato per primo ero io, in altri tu, in altri nessuno dei due. Adesso veniamo interrotti e chi ha chiamato per primo sono io. Per te richiamare sarebbe fare qualcosa di analogo - in base a una analogia - a ciò che ha avuto un buon esito in alcuni dei casi precedenti. Ma noi possiamo ignorare quella analogia, perché in base a essa i precedenti non concordano.

Una volta che siano disponibili diversi precedenti, senza disaccordo o ambiguità sostanziali, non è più necessario che siamo tutti a conoscenza esattamente degli stessi. È sufficiente che ciascuno di noi sia a conoscenza di alcuni precedenti concordanti, che ciascuno si aspetti che ogni altro sia a conoscenza di alcuni precedenti che concordano con i suoi, ecc. È facile rendersi conto del modo in cui ciò potrebbe avvenire: se uno ha incontrato spesso casi in cui, conformandosi a una certa regolarità, veniva conseguita la coordinazione relativamente a un certo problema, e ha incontrato raramente, o mai, casi in cui la coordinazione non veniva conseguita, egli può legittimamente aspettarsi che i suoi vicini abbiano avuto pressapoco la stessa esperienza. Se io ho guidato un po' ovunque negli Stati Uniti e ho visto molti guidare a destra e mai nessuno a sinistra, posso ragionevolmente dedurre che quasi tutti negli Stati Uniti guidano a destra, e che quindi anche colui che sta guidando davanti a me ha visto per la massima parte persone che guidavano a destra, anche se io e lui non abbiamo mai visto la *stessa* persona guidare a destra.

La nostra conoscenza di un precedente non deve necessariamente essere molto dettagliata. È sufficiente sapere che uno conosce molti casi in cui, conformandosi a una certa regolarità, è stata conseguita la coordinazione relativamente a un certo problema. Non c'è alcun bisogno di essere in grado di specificare il tempo e il luogo, gli agenti intervenuti, o qualunque altro particolare; nessun bisogno di essere in grado di ricordare i casi precedenti

uno per uno. Io non sono in grado di citare uno per uno i precedenti in cui gli abitanti degli Stati Uniti hanno guidato a destra; non sono sicuro neppure di saper citare un caso; ciò nonostante, so molto bene di aver visto spesso guidare delle macchine negli Stati Uniti, e quasi sempre erano a destra. E siccome non ho motivo di pensare di aver incontrato un campione anormale, ne deduco che i guidatori degli Stati Uniti guidano quasi sempre a destra; così qualunque guidatore che incontrerò negli Stati Uniti avrà esattamente la mia stessa credenza, si aspetterà che io l'abbia, e così via.

La coordinazione in base al precedente, nella sua accezione più semplice, è questo: conseguimento di coordinazione per mezzo della diretta conoscenza comune del conseguimento di coordinazione in un singolo caso passato, esattamente simile al nostro attuale problema di coordinazione. Eliminando le restrizioni inessenziali, siamo arrivati a questo: conseguimento di coordinazione per mezzo della diretta conoscenza comune di una *regolarità*, la quale governa il conseguimento di coordinazione in una classe di casi passati che hanno una qualche analogia rilevante l'uno con l'altro e con il nostro problema di coordinazione presente. La nostra conoscenza di questa regolarità deriva dalla nostra esperienza relativa ad alcune delle sue manifestazioni, che non sono necessariamente le stesse per tutti.

Data una regolarità nei casi passati, possiamo ragionevolmente dedurre la presenza nel (prossimo) futuro. Siamo infatti autorizzati ad aspettarci che, quando si trovano di fronte a un nuovo problema di coordinazione analogo, gli agenti che sono direttamente a conoscenza di una regolarità passata riusciranno a conseguire la coordinazione seguendo il precedente e continuando a conformarsi alla stessa regolarità. Giungiamo ad aspettarci azioni conformi non solo nei casi passati ma anche in quelli futuri. Acquisiamo la credenza generale, temporalmente illimitata, che, al fine di conseguire la coordinazione, i membri di una data popolazione si conformino a una certa regolarità in un determinato tipo di problema di coordinazione ricorrente.

Ogni nuova azione conforme alla regolarità si aggiunge alla nostra esperienza della conformità generale. La nostra esperienza della conformità generale nel passato ci induce, in virtù dell'efficacia del precedente, a aspettarci una conformità simile nel futuro. E la nostra aspettativa di una conformità futura è un motivo per continuare a conformarci, poiché conformarsi quando gli altri lo fanno significa raggiungere un equilibrio di coordinazione e soddisfare le proprie preferenze. E così via: siamo qui perché siamo qui perché siamo qui perché siamo qui. Una volta che il processo è partito, abbiamo un sistema meta-stabile e autoperpetuantesi di preferenze, aspettative e azioni capaci di persistere indefinitamente. Finché la conformità uniforme è un equilibrio di coordinazione, in modo che ciascuno desideri conformar-

si subordinatamente alla conformità da parte degli altri, l'azione conforme produce l'aspettativa di un'azione conforme e l'attesa di un'azione conforme produce l'azione conforme.

Questo è il fenomeno che io chiamo, *convenzione*. La nostra prima, rozza definizione è:

Una regolarità  $R$  del comportamento dei membri di una popolazione  $P$ , quando essi sono agenti in una situazione ricorrente  $S$ , è una *convenzione* se e solo se, in qualunque esempio di  $S$  tra membri di  $P$ ,

- (1) tutti si conformano a  $R$ ;
- (2) tutti si aspettano che tutti gli altri si conformino a  $R$ ;
- (3) tutti preferiscono conformarsi a  $R$  a condizione che lo facciano gli altri, in quanto  $S$  è un problema di coordinazione e la conformità uniforme a  $R$  è un equilibrio di coordinazione proprio in  $S$ .

## 1.5 Convenzioni campione

Il capitolo 2 sarà dedicato al perfezionamento della definizione. Ma prima di nascondere il concetto dietro i suoi raffinamenti, vediamo come si applica a esempi. Consideriamo alcune convenzioni intese a risolvere i nostri problemi di coordinazione campione.

(1) Se tu e io dobbiamo incontrarci ogni settimana, probabilmente all'inizio ci daremo ogni volta un nuovo appuntamento. Ma dopo che ci siamo incontrati nello stesso posto e alla stessa ora per alcune settimane di seguito, uno di noi dirà alla fine di ogni incontro: "Arrivederci qui la settimana prossima". In seguito non diremo più niente (a meno che la settimana successiva la nostra intesa abituale si riveli insoddisfacente). Entrambi ci recheremo regolarmente in un certo luogo a una data ora ogni settimana, ciascuno andando là per incontrare l'altro, sicuro che questo comparirà. Questa regolarità che si è gradualmente sviluppata nel nostro comportamento è una convenzione.

In questo caso la convenzione che stabilisce il nostro luogo di incontro vale nella più piccola popolazione possibile: soltanto due persone. In altri casi, popolazioni più ampie eventualmente con membri che cambiano hanno luoghi di incontro convenzionali. Ciò che porta in auge un caffè o un bar è l'esistenza, in una cerchia sociale, di una convenzione secondo la quale quello

è il luogo dove andare quando si vogliono avere rapporti sociali. L'uomo di quella canzone che dice: "Stando a un angolo con un dollaro in mano / cerco una donna che cerca un uomo" si trova in *quell'*angolo in conformità a una convenzione che vale tra tutte le prostitute locali e i loro frequentatori.

(2) Nella mia città natale di Oberlin, nell'Ohio, fino a poco tempo fa tutte le chiamate telefoniche locali venivano interrotte senza preavviso dopo tre minuti. Poco tempo dopo che questa abitudine era invalsa, tra gli abitanti di Oberlin si sviluppò una convenzione secondo la quale quando una telefonata veniva interrotta colui che aveva chiamato doveva richiamare, mentre colui che era stato chiamato aspettava. Di solito gli abitanti si conformavano a questa regolarità, attendendosi una conformità da parte dell'altro partecipante alla telefonata. In questo modo la comunicazione veniva facilmente ristabilita, con vantaggio di tutti gli interessati. I nuovi abitanti venivano informati della convenzione o la imparavano per esperienza. La convenzione durò per circa un decennio, finché l'interruzione venne abolita.

Altre regolarità avrebbero potuto dare risultati quasi altrettanto buoni. Avrebbe potuto essere l'utente chiamato a richiamare sempre, oppure quello che precedeva alfabeticamente, oppure il più vecchio. Una qualunque di queste regolarità avrebbe potuto diventare la convenzione operante se un numero sufficiente di noi avesse cominciato a conformarsi. Sarebbe stata un po' meno conveniente della nostra effettiva convenzione; se colui che ha chiamato richiama, egli può ancora ricordare il numero, o perlomeno deve sapere dove trovarlo. Ma gli inconvenienti di un'altra convenzione non sarebbero stati superiori al vantaggio di raggiungere un equilibrio di coordinazione richiamando se e solo se l'altro non richiama.

Questo esempio illustra la possibilità che (descrivendo le azioni in un qualunque modo naturale) una regolarità convenzionale specifichi azioni diverse in condizioni diverse. In questo caso essa specifica quelle che chiameremmo naturalmente azioni diverse per agenti implicati nella situazione S con *ruoli* diversi. Eccettuate descrizioni *ad hoc* come "azione conforme alla regolarità così-e-così", non è necessario che le azioni conformi a una regolarità convenzionale condividano una qualunque descrizione naturale comune. Perciò, quando parliamo della convenzione di compiere un'azione A in una situazione S, si deve intendere che A può stare per una descrizione d'azione innaturalmente complessa.

(3) Se riescono in qualche modo a prendere un ritmo uniforme e lo mantengono per un po' di tempo, i due rematori della barca di Hume "fanno ciò in virtù di un accordo o convenzione, anche se non si sono mai scambiati una promessa". Una regolarità di comportamento - il remare con quel particolare ritmo - persiste perché gli interessati si aspettano che venga protratta e

perché essi vogliono accordare i loro ritmi di voga. “Questa consapevolezza dell’interesse comune... condivisa da entrambi... produce una risoluzione e un comportamento adeguato” in cui “le azioni di ciascuno... sono in rapporto con quelle altrui, e le compiamo in base alla supposizione che l’altro dovrà compierne certe altre.”

Questa convenzione è singolare. Vale in una popolazione molto piccola per un tempo molto breve - tra due persone, per alcuni minuti -, e la regolarità è di un genere che avremmo molte difficoltà a descrivere, anche se siamo facilmente in grado di afferrarla. Ma queste caratteristiche singolari non compromettono la sua convenzionalità.

(4) Guidiamo sulla corsia destra delle strade degli Stati Uniti (o sulla corsia sinistra delle strade dell’Inghilterra, dell’Australia, della Svezia prima del 1967, di parte dell’Austria prima di una certa data, e di altri luoghi) perché non vogliamo guidare sulla stessa corsia di coloro che vengono verso di noi, e ci aspettiamo che essi guidino a destra.

C’è una complicazione: se non guidiamo a destra, la pattuglia della polizia stradale ci raggiunge e noi veniamo puniti. Abbiamo quindi un incentivo indipendente per guidare a destra, e questo incentivo è indipendente da come guidano gli altri. Ma non si tratta di una differenza importante. Se io mi aspettassi la guida a sinistra da parte degli altri, anch’io mi metterei a sinistra, pattuglia o no. La mia preferenza a guidare sulla stessa corsia degli altri è superiore a qualunque incentivo la pattuglia possa fornirmi per guidare a destra. E lo stesso vale per quasi ogni altro, ne sono sicuro. La pattuglia modifica le vincite in favore della guida a destra; ma ci sono ancora due equilibri di coordinazione diversi. Le punizioni sono superflue se concordano con la nostra convenzione, vengono superate in efficacia se contrastano con essa, ma in entrambi i casi non sono decisive, e quindi non rendono affatto meno convenzionale la guida a destra. Lo stesso vale per altre considerazioni che favoriscono un equilibrio di coordinazione rispetto all’altro: il fatto che le automobili abbiano la guida a sinistra, il fatto che la maggior parte di noi usi di preferenza la destra, e così via.

(5) Se quattro persone che campeggiano insieme scoprono che, perlustrando lo stesso terreno in cerca di legna da ardere, sprecano energie, può darsi che si stanchino di ciò e si mettano d’accordo una volta per tutte: Morgan cercherà a nord, Jones a est, Owen a sud e Griffith a ovest. Da quel giorno in poi ciascuno va per la propria strada senza ulteriori discussioni. Una regolarità è cominciata per accordo esplicito. All’inizio, forse, essa dura perché ciascuno si sente vincolato dalla promessa e non fa alcun conto dei vantaggi del mantenerla o dell’infrangerla. Ma gli anni passano. Gli interessati dimenticano di essersi messi d’accordo. Morgan è sostituito da Thomas,

che non ha mai saputo nulla dell'accordo e non ha mai promesso alcunché. Eppure, ogniqualvolta c'è bisogno di legna da ardere, ciascuno se ne va nella propria direzione, perché sa che in quel modo avrà il terreno tutto per sé. Nella misura in cui il potere della loro promessa iniziale va diminuendo, la regolarità del loro comportamento diventa una convenzione.

(6) Desiderando frequentare dei party vestiti come saranno vestiti gli altri, noi indossiamo, per ogni occasione, l'abito convenzionale, qualunque esso sia; scegliendo i vestiti noi agiamo in conformità a una convenzione della nostra cerchia sociale. Per mezzo di una regolarità convenzionale condizionale, la quale specifica lo stile degli abiti indossati in varie circostanze, noi soddisfacciamo il nostro interesse comune a vestirci in modo simile.

Ma dobbiamo distinguere due casi. Se ciascuno di noi vuole vestirsi come la maggioranza e vuole che anche ciascun altro si vesta come la maggioranza, allora raggiungiamo un equilibrio di coordinazione quando ci vestiamo tutti in modo simile: la nostra regolarità è una convenzione autentica. Supponiamo invece che molti di noi siano persone dispettose che vogliono sì vestirsi come la maggioranza, ma che vogliono anche che ci sia una minoranza vestita diversamente, per deriderla. Quando ci vestiamo tutti in modo simile, raggiungiamo ancora un equilibrio, ma non un equilibrio di coordinazione: nessuno vorrebbe essersi vestito diversamente, ma i dispettosi vorrebbero che alcune altre persone - per esempio, i peggiori nemici - si fossero vestite diversamente. La regolarità per mezzo della quale raggiungiamo questo equilibrio non è, in base alla mia definizione, una convenzione autentica, in quanto l'elemento del conflitto di interesse le impedisce di essere un mezzo per conseguire un equilibrio di *coordinazione*.

Può darsi che non sia ovvio che le nostre regolarità di abbigliamento non debbano essere chiamate convenzioni, e questo nel caso che ci siano molte persone che desiderano vederle violate. Ma dato che la nostra analisi ci ha mostrato come la presenza di un conflitto effettivo introduca una diversità tra questo caso e gli altri chiari casi di convenzione, e introduca un'importante analogia tra questo caso e gli altri chiari casi di non-convenzione come quelli esaminati nel capitolo 3.5, penso che la nostra conclusione debba essere in accordo con l'analisi, anche se questa contrasta con la nostra prima impressione. Se il lettore non è d'accordo, posso soltanto ricordargli che l'analisi da me intrapresa non riguarda il concetto di convenzione elaborato da qualcun altro, ma solo il mio.

(7) Se siamo oligopolisti arrivati che intendono mantenere, per le loro merci, un prezzo uniforme ma fluttuante, non osiamo stipulare un qualunque accordo esplicito sui prezzi, in quanto si tratterebbe di una congiura per il controllo del commercio, ma possiamo giungere a un'intesa tacita -

cioè a una convenzione - in base ai nostri modi di rispondere ciascuno ai prezzi dell'altro. Potremmo per esempio cominciare con il seguire qualcuno che imponga i prezzi: una persona risoluta, che prende l'iniziativa di modificare i prezzi, opportunamente attento a fissarli entro una gamma che sia soddisfacente per tutti noi.

In questo esempio diventa seriamente artificioso segmentare la nostra attività continua in una serie di problemi di coordinazione separati e analoghi, correlati in virtù della sola efficacia del precedente. La difficoltà si ripresenterà negli esempi (9), (10) e (11); era presente, in qualche misura, in (3) e (4).] Noi possiamo effettivamente fissare o riprendere in considerazione i prezzi ogni volta. Quanto è lungo un problema di coordinazione? Proviamo a fingere, compiendo già un'idealizzazione, di fissare i nostri prezzi ogni mattina e di non poterli modificare durante la giornata. In tal caso ogni giorno lavorativo è un problema di coordinazione. Ma un giorno è troppo breve. I nostri clienti impiegano più di un giorno a fare il giro dei negozi; essi confrontano il mio prezzo di oggi con il tuo di ieri e con quello di domani di qualcun altro. Stiamo trascurando gran parte della coordinazione: la coordinazione dell'azione di uno in un giorno con l'azione di un altro in un altro giorno immediatamente precedente o successivo. Se, d'altra parte, consideriamo come problemi di coordinazione intervalli di tempo più lunghi, allora - contrariamente alla definizione - ciascuno ha tempo di fare più scelte diverse entro un singolo problema di coordinazione. Potremmo fingere che ciascuno parta ogni settimana scegliendo un piano circostanziale che specifichi la condotta da seguire durante la settimana in ogni possibile circostanza (una *strategia* nel senso della teoria dei giochi), e poi segua per tutta la settimana il suo piano senza fare scelte ulteriori. In questo caso una settimana lavorativa è un problema di coordinazione in cui ciascuno fa soltanto la scelta iniziale del piano circostanziale. Ma questa impostazione del problema descrive assai male quello che facciamo; inoltre, trascura di nuovo la coordinazione, per esempio, tra i miei prezzi di venerdì e i tuoi del lunedì successivo. Un rimedio migliore, registrato qui in vista del capitolo II.3, va in profondità. Possiamo trascurare i problemi di coordinazione individuali; invece di dire che la conformità uniforme alla regolarità  $R$  costituisce un equilibrio di coordinazione in ogni esempio della situazione  $S$ , possiamo dire approssimativamente la stessa cosa in termini di preferenze condizionali per la conformità ad  $R$ .

(8) Se i cacciatori di cervi di Rousseau partecipano ogni volta alla caccia, lo fanno per convenzione. Ciascuno partecipa perché spera che gli altri partecipino, come hanno fatto in precedenza, ed egli mangerà meglio partecipando e prendendosi la sua parte del cervo quando questo verrà catturato.

Ma, in modo meno ovvio, se essi si separano sempre per cacciare conigli ciascuno per conto suo, anche questa è una convenzione. Se la caccia al cervo fallisce a meno che vi prendano parte tutti, non c'è alcuno scopo a unirsi, a meno che lo facciano anche tutti gli altri. Ciascuno preferisce cacciare conigli se anche uno solo degli altri lo fa, e *a fortiori* se tutti gli altri lo fanno. Per ciascuno, cacciare il proprio coniglio non è un buon equilibrio di coordinazione. Ma nondimeno è un equilibrio di coordinazione, se è vero che catturare un coniglio è sempre meglio che partire per una caccia solitaria al cervo, destinata a fallire. Così la cattura del coniglio è, per definizione, una convenzione.

(9) Nell'ipotesi che ciascuno di noi desideri lo sfruttamento esclusivo di qualche terreno e che nessuno pensi che valga la pena di cercare di rilevare da un altro lo sfruttamento di qualche appezzamento di terreno, ogni divisione *de facto* della terra è una convenzione. Ciascuno continua a sfruttare una certa porzione tenendosi lontano dagli altri terreni nella consapevolezza che, siccome gli altri continueranno a sfruttare tutti gli altri terreni, questo è l'unico modo di soddisfare i propri bisogni e di tenersi lontani dai guai. Una convenzione migliore potrebbe fornire un vero e proprio criterio che regolasse gli avvicendamenti nella popolazione dei coltivatori della terra. Potrebbe far parte della convenzione, per esempio, che quando qualcuno muore, il più anziano dei giovani che non lavorano ancora la terra cominci lo sfruttamento della porzione vacante.

Non ho chiamato queste porzioni di terra *proprietà*. Nel caso più semplice per esempio, in una società anarchica l'istituzione della proprietà potrebbe essere nient'altro che una convenzione che specificasse a chi sia riservato l'uso esclusivo di certi beni. Questa sembra essere la teoria della proprietà di Hume. Per noi, l'istituzione della proprietà è più complicata; l'abbiamo inserita in un elaborato sistema di leggi e istituzioni. Noi non diciamo che uno *squatter*<sup>6</sup> possiede la terra che coltiva, sebbene egli goda per convenzione dello sfruttamento esclusivo di essa; questo perché un altro richiedente è autorizzato dalla legge a chiamare la polizia per far cacciar via lo squatter. Per questo io non *definirò* la proprietà come bene riservato per convenzione all'uso esclusivo da parte di qualcuno.

(10) Un mezzo di scambio - per esempio la moneta - ha le sue prerogative particolari in virtù della convenzione tra i commercianti di accettarlo senza obiezioni in cambio di merci e di servizi. Alcuni mezzi convenzionali sono migliori di altri; quelli ingombranti e quelli deperibili non vanno bene; quelli

---

<sup>6</sup>Negli Stati Uniti, persona che occupa un terreno per ottenere dal governo il diritto di conservarlo. (N.d.T.)

che conserverebbero una certa utilità se la convenzione venisse meno vanno bene; ma lo svantaggio dell'accettare un mezzo di scambio cattivo è minore dello svantaggio di rifiutarlo quando gli altri lo accettano, o di accettare qualcosa che non si può usare né spendere. Di nuovo, come nel caso (4), c'è la complicazione delle sanzioni legali. Il rifiuto di accettare valuta legale rende un debito legalmente inesigibile. Ma, di nuovo, tali sanzioni sono superflue se concordano con la convenzione, vengono superate in efficacia se contrastano con essa: in entrambi i casi non sono decisive, e quindi non rendono affatto meno convenzionale la nostra regolarità.

Penso che possiamo tranquillamente definire un *mezzo di scambio* come qualunque bene che, in qualche popolazione, è accettato convenzionalmente in cambio di merci e servizi. Questa definizione fa sorgere un problema imbarazzante: è giusto dire che abbiamo la convenzione di accettare i nostri mezzi di scambio come compenso per merci e servizi? È falso dire che la nostra convenzione è che noi accettiamo i nostri mezzi di scambio come compenso per merci e servizi. Infatti, ciò che segue "che" non esprime alcuna convenzione perché è vero, per definizione, di qualunque popolazione. D'altra parte, è vero dire *dei* nostri mezzi di scambio che la nostra convenzione è di accettarli come compenso per merci e servizi. La mia domanda era ambigua. Può essere letta in modo opaco o in modo trasparente.<sup>7</sup> È come chiedersi se Hegel sapeva che il numero dei pianeti è maggiore di sette. Egli non sapeva che il numero dei pianeti è maggiore di sette. Ma sapeva, *del* numero dei pianeti - cioè del nove - che *esso* è maggiore di sette.

(11) L'uso comune da parte di una popolazione di una qualche lingua il gallese, per esempio è una convenzione. I gallesi di varie parti del Galles usano il gallese; ciascuno usa il gallese perché si aspetta che lo usino i suoi vicini, e perché, per amore della comunicazione, egli desidera usare qualunque lingua usino i suoi vicini.

Non è vero piuttosto che egli desidera usare qualunque lingua i suoi vicini *capiscano*? Sì; ma, per un fenomeno caratteristico della natura umana, egli e i suoi vicini comprenderanno nel modo migliore la lingua che usano. Così, il modo esatto di esprimersi è che egli desidera usare la lingua che essi usano perché è quella la lingua che essi capiranno. Ne segue che questo è un altro caso di coordinazione nel tempo: egli desidera usare la lingua che essi hanno usato maggiormente in un periodo del passato, un periodo abbastanza lungo perché essi si siano impadroniti dell'uso della lingua.

Dire che egli desidera usare qualunque lingua i suoi vicini usino non equi-

---

<sup>7</sup>Cfr. W.V.O. Quine, "Quantifiers and Propositional Attitudes", *Journal of Philosophy*, 53 (1956), pp. 177-187.

vale a dire che se essi, in qualche modo, cambiassero improvvisamente, egli desidererebbe cambiare immediatamente. Non lo desidererebbe, perché non potrebbe; dovrebbe esercitarsi nella loro nuova lingua. D'altronde, potrebbe contare sulla loro comprensione del gallese per un certo periodo dopo l'abbandono di esso. Ma probabilmente egli desidererebbe cambiare il più presto possibile. E se improvvisamente accadesse che i suoi vicini stanno usando la nuova lingua da vent'anni - mentre egli, per esempio, dormiva come Rip Van Winkle - il nostro gallese cercherebbe di conformarsi con la massima urgenza.

Io non nego, naturalmente, che una persona possa preferire una lingua a un'altra: per esempio, la lingua dei suoi padri a quella dei loro conquistatori. Ma questo non importa. Equilibri di coordinazione diversi non devono essere ugualmente buoni, ma solo buoni abbastanza perché tutti siano pronti a svolgere la propria parte se la svolgono gli altri. Sono pochi quelli che rinuncerebbero a comunicare per amore della lingua materna, se si giungesse a tanto.

Certamente non ogni carattere di una lingua è convenzionale. Nessuna lingua umanamente possibile fa uso di sibili ultrasonici, per cui non è per convenzione che i gallesi non ne fanno uso. Non sappiamo ancora con esattezza quali caratteri delle lingue sono convenzionali e quali sono comuni a tutte le lingue umanamente possibili; Noam Chomsky e la sua Scuola hanno sostenuto che la convenzionalità è minore di quanto si sarebbe potuto pensare.<sup>8</sup> Ma, finché almeno due lingue sono umanamente possibili, deve essere per convenzione che una comunità sceglie di usare l'una o l'altra.

Dicendo che i gallesi usano il gallese per convenzione io non affermo che è una convenzione che i gallesi usino il gallese. Questo, o qualcosa di simile ma più complicato, potrebbe forse essere vero per definizione di "gallese". Io invece dico *del* gallese, che il fatto che *essi lo* usino è una convenzione tra i gallesi. La differenza è la stessa ambiguità tra lettura opaca e lettura trasparente che si era presentata in (10).

Se è vero che l'uso del gallese è una convenzione, deve essere una regolarità di comportamento. Non si tratta, naturalmente, di una regolarità che determina completamente il comportamento di un gallese. Egli può dire un gran numero di cose, o rimanere in silenzio, e può rispondere ad emissioni verbali in un gran numero di modi, continuando a conformarsi alla regolarità convenzionale. Ma in ciò non c'è nulla di speciale. Nessuna convenzione

---

<sup>8</sup>Noam Chomsky, "Recent Contributions to the Theory of Innate Ideas", *Synthese*, 17 (1967), pp. 2-11. [Trad. it. di A. De Palma, in N. Chomsky, *Saggi linguistici*, Boringhieri, Torino, 1969, Vol. III. (N.d.T.)]

determina tutte le particolarità del comportamento. (La convenzione del luogo d'incontro, per esempio, non specifica se bisogna recarsi nel luogo d'incontro a piedi o a cavallo.) Questa convenzione, come qualunque altra, limita il comportamento senza sopprimere tutte le scelte. Ci sono più scelte, e più scelte importanti, in questo caso che in alcuni altri; ma non c'è alcuna differenza di specie.

Una convenzione è una regolarità di comportamento. Non stendo dire che gli utenti del gallese si conformano alla loro convenzione quando e solo quando di essi si può dire correttamente che stanno "usando il gallese". Un uomo che mente in gallese sta usando il gallese, ma viola la sua convenzione; un uomo che rimane silenzioso durante una conversazione può darsi che si conformi alla convenzione anche se non usa il gallese. A tempo debito vedremo come può essere descritta la convenzione di una lingua; qui dirò soltanto che è una regolarità che condiziona la produzione di emissioni verbali e di iscrizioni e la risposta a esse. La competenza linguistica consiste in parte nella disposizione a conformarsi con facilità a quel condizionamento; e in parte nell'aspettativa che i propri vicini avranno disposizioni analoghe, con il riconoscimento della loro conformità come motivo della propria. Senza dubbio un bambino o un idiota può conformarsi senza motivo; se è così, egli non partecipa della convenzione e la sua competenza linguistica è incompleta.



## Capitolo 2

# Raffinamento della convenzione

### 2.1 Conoscenza comune

L'accordo, la salienza, o il precedente, come abbiamo visto, possono risolvere un problema di coordinazione producendo un sistema di aspettative reciproche concordanti di primo ordine e di ordine superiore. Basta che immaginiamo degli esempi per convincerci che aspettative di ordine superiore *verrebbero* prodotte. Ma come? Quali premesse ci autorizzano a concludere che gli altri hanno certe aspettative, che gli altri si aspettano che gli altri abbiano certe aspettative, e così via? E come avviene che il processo si interrompa - il che sicuramente avviene in modo da produrre soltanto le aspettative degli ordini più bassi?

Prendiamo un caso semplice di coordinazione per accordo. Supponiamo che valga il seguente stato di cose - lo chiameremo A: tu e io ci siamo incontrati, ci siamo parlati, e tu devi andartene prima che il nostro affare sia concluso; perciò dici che tornerai nello stesso posto domani. Immaginiamo ciò che succede. Chiaramente, io mi aspetterò che tu ritorni. Tu ti aspetterai che io mi aspetti che tu ritorni. Io mi aspetterò che tu ti aspetti che io mi aspetti che tu ritorni. Forse ci saranno ancora uno o due ordini. Che cosa c'è in A che spiega la generazione di queste aspettative di ordine superiore? Io suggerisco che la ragione sia che A soddisfa queste tre condizioni:

- (1) Tu e io abbiamo motivo di credere che A valga.
- (2) A indica a entrambi che tu e io abbiamo motivo di credere che A valga.

(3) A indica a entrambi che tu tornerai.

Che cos'è l'*indicare*? Diciamo che A *indica* a qualcuno, x, che se e solo se, se x avesse motivo di credere che A valesse, x avrebbe perciò motivo di credere che—. Ciò che A indica a x dipenderà quindi dai modelli induttivi e dalle informazioni di base di x.

Le tre premesse principali (1), (2) e (3), unitamente a pre-messe ausiliarie relative alla nostra razionalità, ai nostri modelli induttivi e alle informazioni di base, saranno sufficienti a giustificare le mie aspettative di ordine superiore. Vediamo come procederebbe il mio ragionamento.

Si osservi che se A indica qualcosa a x, e se y ha gli stessi modelli induttivi e la stessa informazione di base di x, allora A deve indicare la stessa cosa a y. Quindi, se A indica a x che y ha motivo di credere che A valga, e se A indica a x che—, e se x ha motivo di credere che y disponga dei suoi stessi modelli induttivi e delle sue stesse informazioni di base, allora A indica a x che y ha motivo di credere che (dove tale motivo è il motivo che y ha di credere che A valga). Supponiamo che tu e io abbiamo motivo di credere che disponiamo degli stessi modelli induttivi e delle stesse informazioni di base, perlomeno in misura sufficiente perché A indichi la stessa cosa per entrambi. Allora (2), applicato a (3), implica:

(4) A indica a entrambi che ciascuno di noi ha motivo di credere che tu tornerai.

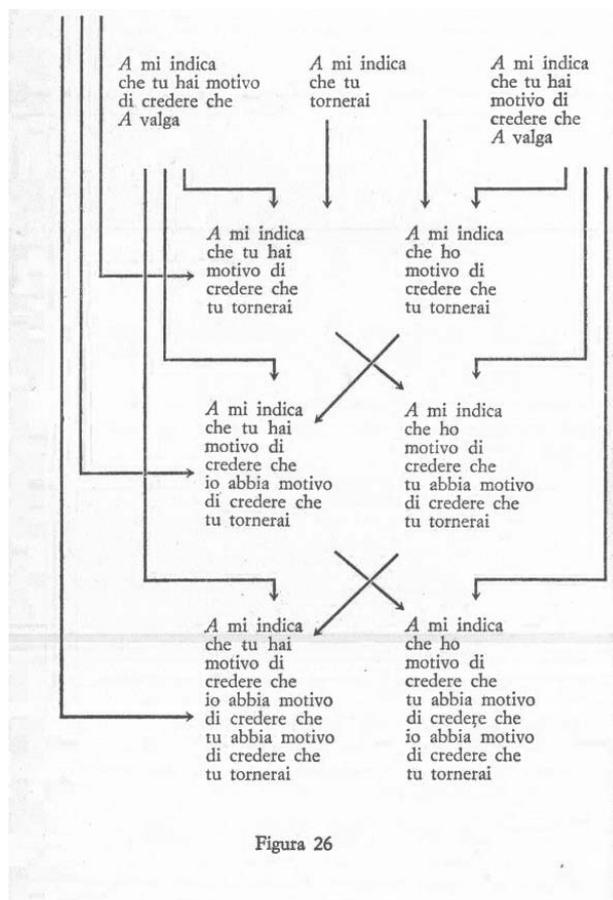
E (2), applicato a sua volta a (4), implica:

(5) A indica a entrambi che ciascuno di noi ha motivo di credere che l'altro ha motivo di credere che tu tornerai.

E così via *ad infinitum*, in quanto ogni nuova conclusione comincia con "A indica a entrambi che..." Si osservi che si tratta di una catena di implicazioni, non di passaggi del ragionamento effettivo di qualcuno. Non c'è quindi nulla di improprio nella sua lunghezza infinita. La Figura 26 è una rappresentazione più dettagliata di tali implicazioni nel mio caso; nel tuo caso, le implicazioni possono venir rappresentate in modo analogo.

Si osservi inoltre che la nostra definizione di indicazione fornisce un principio di separazione: se A indica a x che e x ha motivo di credere che A valga, allora x ha motivo di credere che—. La premessa (1), applicata in questo modo a (3), implica:

(3') Ciascuno di noi ha motivo di credere che tu tornerai.



La premessa (1), applicata a (4), implica:

- (4') Ciascuno di noi ha motivo di credere che l'altro abbia motivo di credere che tu tornerai.

La premessa (1) applicata a (5) implica:

- (5') Ciascuno di noi ha motivo di credere che l'altro abbia motivo di credere che il primo abbia motivo di credere che tu tornerai.

E così via, per tutta la sequenza infinita che abbiamo considerato sopra. Io non sto ancora parlando dell'effettivo ragionamento di qualcuno o di ciò che qualcuno crede effettivamente. Ma l'unico ragionamento effettivo di cui ora c'è bisogno è un ragionamento che converti queste iterazioni di "ha motivo di credere" nelle corrispondenti iterazioni di "crede". A questo fine abbiamo bisogno di premesse ausiliari relative alla razionalità.

Chiunque abbia motivo di credere qualcosa passerà a credere quella cosa, a condizione che abbia un grado sufficiente di razionalità. Così, in virtù di (3'), se abbiamo entrambi un grado sufficiente di razionalità, allora si verificherà che

- (3'') Ciascuno di noi si aspetta che tu tornerai.

In virtù di (4'), se ciascuno di noi ha motivo di attribuire un grado sufficiente di razionalità all'altro, allora ciascuno ha motivo di aspettarsi che l'altro si aspetti che tu tornerai. Se, inoltre, abbiamo entrambi un grado sufficiente di razionalità, allora si verificherà che

- (4'') Ciascuno di noi si aspetta che l'altro si aspetti che tu tornerai.

In virtù di (5'), se ciascuno di noi ha motivo di aspettarsi che l'altro abbia motivo di attribuirgli un grado sufficiente di razionalità, allora ciascuno ha motivo di aspettarsi che l'altro abbia motivo di aspettarsi che egli si aspetti che tu tornerai. Se, inoltre, ciascuno di noi ha motivo di attribuire all'altro un grado sufficiente di razionalità, allora ciascuno ha motivo di aspettarsi che l'altro si aspetti che egli si aspetti che tu tornerai. E se, inoltre, abbiamo entrambi un grado sufficiente di razionalità, si verificherà che

- (5'') Ciascuno di noi si aspetta che l'altro si aspetti che egli si aspetti che tu tornerai.

E così via. Ciascun termine della sequenza (3'), (4'), (5') ..., assieme alla razionalità sufficiente, al motivo di attribuire una razionalità sufficiente, ecc.,

garantisce il costituirsi della corrispondente aspettativa di primo ordine o di ordine superiore. Ma i gradi di razionalità che ci è richiesto di avere, di aver motivo di attribuire, ecc., crescono naturalmente in modo rapido. È per questo motivo che si costituiscono effettivamente soltanto le aspettative degli ordini più bassi. Il processo di generazione si ferma quando vengono meno le premesse ausiliarie.

Abbiamo così completato la nostra illustrazione di uno stato di cose che produce aspettative di ordine superiore. Penso che questo esempio sia tipico; tutte le aspettative di ordine superiore che contribuiscono a consolidare le convenzioni, e più' o meno tutte le aspettative che abbiamo in assoluto, sembrano prodotte in questo modo.

Diciamo che in una popolazione P è conoscenza comune che\_\_ se e solo se vale un certo stato di cose A tale che:

- (1) Tutti in P hanno motivo di credere che A valga.
- (2) A indica a tutti in P che tutti in P hanno motivo di credere che A valga.
- (3) A indica a tutti in P che\_\_

Possiamo chiamare un tale stato di cose A base della conoscenza comune in P che\_\_. A fornisce ai membri di P parte di ciò di cui essi abbisognano per costituire aspettative di ordine arbitrario che\_\_, relative a sequenze di membri di P. La parte che essa fornisce loro è la parte peculiare al contenuto \_\_. Il rimanente di cui abbisognano è ciò di cui abbisognano per costituire *qualunque* aspettativa di ordine superiore nel modo che stiamo considerando: attribuzione reciproca di certi modelli induttivi comuni e di certe informazioni di base, razionalità, attribuzione reciproca di razionalità, e così via.

Torniamo al nostro esempio e consideriamo in modo più completo lo stato di cose A. Supponiamo che, come parte di A, noi manifestiamo le nostre preferenze condizionali a ritornare nel luogo d'incontro. Allora A può indicarci anche che abbiamo entrambi tali preferenze. Se ciò avviene, A può servire come base non solo della conoscenza comune che tu tornerai, ma anche come base della conoscenza comune che ciascuno di noi preferisce tornare se l'altro torna. Supponiamo anche che, come parte di A, noi manifestiamo in qualche modo un minimo di razionalità. Allora A può indicarci il nostro possesso di questo minimo di razionalità, ed essere la base della conoscenza comune di esso. Ma ora A - l'evento del nostro accordarci di ritornare - genera tutte le aspettative di ordine superiore che contribuiscono al nostro successo nella risoluzione del problema di coordinazione per ricostruzione.

Una base della conoscenza comune genera aspettative di ordine superiore con l'aiuto di preesistenti aspettative di razionalità di ordine superiore. Possono queste stesse aspettative essere generate da qualche base della conoscenza comune? Sì, perché tutte le aspettative di razionalità di ordine superiore che sono necessarie per generare un'aspettativa di n-esimo ordine sono di ordine inferiore a n. Ciò che interrompe la generazione di aspettative di ordine superiore è la limitata quantità di razionalità indicata da qualunque base, e non una difficoltà nella generazione di aspettative di ordine superiore di altrettanta razionalità quanta è indicata da una base.

L'accordo di svolgere la propria parte in un equilibrio di coordinazione è una base della conoscenza comune che tutti svolgeranno la loro parte. La salienza è un'altra base della conoscenza comune che tutti svolgeranno la loro parte in un equilibrio di coordinazione; ma si tratta, in generale, di una base più debole che genera aspettative di ordine superiore più deboli, perché la salienza di un equilibrio non è un'indicazione molto forte del fatto che gli agenti tenderanno a sceglierlo. Anche i precedenti sono una base della conoscenza comune che tutti svolgeranno la loro parte in un equilibrio di coordinazione; e, in particolare, la conformità passata a una convenzione è una base della conoscenza comune di una tendenza a continuare a conformarsi. Consideriamo una regolarità convenzionale R in una popolazione P. Tutti in P hanno motivo di credere che i membri di P si siano conformati nel passato a R. Il fatto che i membri di P si siano conformati a R nel passato indica a tutti in P che tutti in P hanno motivo di credere che i membri di P si siano conformati, nel passato, a R. E il fatto che i membri di P si siano conformati a R nel passato indica a tutti in P che essi tenderanno a fare altrettanto anche in futuro.

Per esempio, i guidatori degli Stati Uniti hanno finora guidato a destra. Noi tutti abbiamo motivo di credere che sia così. E il fatto che sia così indica a noi tutti che noi tutti abbiamo motivo di credere che i guidatori degli Stati Uniti abbiano finora guidato a destra e anche che i guidatori degli Stati Uniti tenderanno in avvenire a guidare a destra.

Le nostre condizioni definitorie per l'esistenza di una convenzione consistono di una regolarità di comportamento, di un sistema di aspettative reciproche e di un sistema di preferenze. Propongo di correggere la definizione: non solo queste condizioni devono essere soddisfatte, ma deve anche esserci, nella popolazione, la conoscenza comune che lo sono. La nostra definizione corretta suona così:

Una regolarità R nel comportamento dei membri di una popolazione P, quando essi sono agenti in una situazione ricorrente S,

è una *convenzione* se e solo se è vero che, e in P è conoscenza comune che, in qualunque esempio di S tra membri di P,

- (1) tutti si conformano a R;
- (2) tutti si aspettano che tutti gli altri si conformino a R;
- (3) tutti preferiscono conformarsi a R a condizione che lo facciano gli altri, in quanto S è un problema di coordinazione e la conformità uniforme a R è un equilibrio di coordinazione in S.

Così, deve esserci un qualche stato di cose A (tale che A vale, tutti in P hanno motivo di credere che A valga, e A indica a tutti in P che tutti in P hanno motivo di credere che A valga) che indica a tutti in P che i membri di P si conformano a R, che essi si aspettano reciprocamente di conformarsi a R, e che essi hanno preferenze che fanno della conformità uniforme a R un equilibrio di coordinazione.

Una ragione per correggere la definizione di convenzione è il semplice fatto che vogliamo inserire nella definizione tutte le caratteristiche importanti che sono comuni ai nostri esempi, e la conoscenza comune dei fatti pertinenti sembra essere una caratteristica di questo tipo. C'è un'altra ragione: la correzione ci aiuta ad affrontare certi casi anomali, regolarità che intuitivamente sembrano diverse dai casi chiari di convenzione ma che in base alla definizione non corretta si sarebbero descritte come convenzioni.

Supponiamo che ciascuno guidi a destra perché si aspetta che ciascun altro guidi a destra e intende evitare delle collisioni. Ma supponiamo che nessuno abbia nell'intelligenza degli altri una fiducia uguale a quella che ha nella propria. Ciascuno ha questa falsa credenza (chiamiamola *f*): "Tranne me, tutti guidano a destra per abitudine, per nessun motivo, e continuerebbero a guidare a destra indipendentemente dalle loro aspettative circa gli altri." In base alla definizione non corretta questo è un caso di convenzione, nonostante le credenze false; ma penso che andrebbe escluso. In base alla definizione corretta non può essere un caso di convenzione (a meno che siamo estremamente irrazionali); se infatti lo è, c'è qualche stato di cose che abbiamo motivo di credere che valga e che ci indica che *f* è falsa. Questo esempio è solo il primo di una serie. Supponiamo adesso che nessuno abbia in realtà la falsa credenza *f*, ma che ciascuno la attribuisca falsamente a tutti gli altri. Anche questo, in base alla definizione corretta, non può essere un caso di convenzione (a meno che siamo estremamente irrazionali); se infatti lo è, c'è qualche stato di cose che abbiamo motivo di credere che valga e che

ci indica che ciascuno ha motivo, di non credere a *f*. E così via. Gli esempi diventano sempre più improbabili, ma non meno meritevoli di esclusione; la definizione corretta continua a escluderli (data un'assunzione di razionalità sufficientemente forte - via via che procediamo sono necessarie assunzioni di razionalità sempre più forti).

Si potrebbe ormai supporre che la conoscenza comune sia l'unica fonte possibile di aspettative di ordine superiore. Ma non è così; c'è un metodo generale per la produzione di aspettative di ordine arbitrariamente elevato in isolamento. Per esempio, io posso acquisire un'aspettativa isolata di quarto ordine nel modo seguente. Supponiamo che io sia un cittadino di Ableton e che creda a tutto ciò che viene stampato sull' "Ableton Argus". L' "Argus" di oggi pubblica questo racconto:

Il "Bakerville Bugie" è assolutamente inattendibile; ciò che esso pubblica può essere tanto vero quanto falso. Eppure i cittadini di Bakerville credono a tutto ciò che viene stampato su di esso. Il "Bugie" di oggi pubblicava questo servizio:

Il "Charlie City Crier" è assolutamente inattendibile; ciò che esso pubblica può essere tanto vero quanto falso. Eppure i cittadini di Charlie City credono a tutto ciò che viene stampato su di esso. Il "Crier" di oggi pubblicava questo servizio:

Il "Dogpatch Daily" è assolutamente inattendibile; ciò che esso pubblica può essere tanto vero quanto falso. Eppure i cittadini di Dogpatch credono a tutto ciò che viene stampato su di esso. Il "Daily" di oggi pubblicava questo servizio: Domani pioveranno gatti e topi.

Io non mi aspetterei che piovano gatti e topi. Io non mi aspetterei che i cittadini di Dogpatch si aspettino che piovano gatti e topi. Io non mi aspetterei che i cittadini di Charlie City si aspettino che i cittadini di Dogpatch si aspettino che piovano gatti e topi. Ma mi aspetterei che i cittadini di Bakerville si aspettino che i cittadini di Charlie City si aspettino che i cittadini di Dogpatch si aspettino che piovano gatti e topi. In altre parole, avrei un'aspettativa di quarto ordine che piovano gatti e topi senza alcuna corrispondente aspettativa di ordine inferiore che ciò avverrà. Ovviamente, il metodo avrebbe funzionato per una serie di giornali arbitrariamente

lunga; la serie avrebbe potuto ripetersi, a condizione che due termini adiacenti non fossero lo stesso. Non pretendo che questo metodo di generazione di aspettative di ordine superiore isolate sia di grande importanza pratica; esso stabilisce semplicemente la possibilità di tale generazione.

## 2.2 Conoscenza delle convenzioni

Supponiamo che in una popolazione  $P$  sia conoscenza comune che un qualche stato di cose  $B$  vale. Allora tutti in  $P$  hanno motivo di aspettarsi che in  $P$  sia conoscenza comune che  $B$  vale. Infatti, per la definizione di conoscenza comune, c'è un qualche stato di cose  $A$  tale che:

- (1) tutti in  $P$  hanno motivo di credere che  $A$  valga;
- (2)  $A$  indica, in  $P$ , che tutti in  $P$  hanno motivo di credere che  $A$  valga;
- (3)  $A$  indica, in  $P$ , che  $B$  vale.

Da (1) e (2) possiamo inferire:

- (4) tutti in  $P$  hanno motivo di credere che tutti in  $P$  abbiano motivo di credere che  $A$  valga.

Dal solo (2) possiamo inferire:

- (5) tutti in  $P$  hanno motivo di credere che  $A$  indichi, in  $P$ , che tutti in  $P$  hanno motivo di credere che  $A$  valga.

Analogamente, da (3) possiamo inferire:

- (6) tutti in  $P$  hanno motivo di credere che  $A$  indichi, in  $P$ , che  $B$  vale.

E da (4), (5) e (6) possiamo inferire che tutti in  $P$  hanno motivo di credere che ci sia uno stato di cose  $A$  che soddisfa le condizioni (1), (2) e (3).

Così se una convenzione, in particolare, vale come una conoscenza comune, allora appartenere alla popolazione nella quale quella convenzione vale - essere parte di essa - è sapere, in qualche modo, che la convenzione vale. Se una regolarità  $R$  è una convenzione nella popolazione  $P$ , allora deve essere vero, e deve essere conoscenza comune in  $P$ , che  $R$  soddisfa le condizioni definitorie di una convenzione. Se è conoscenza comune che  $R$  le soddisfa, allora tutti in  $P$  hanno motivo di credere che sia vero, e che sia conoscenza comune in  $P$ , che  $R$  le soddisfa; il che equivale a dire che tutti in  $P$  devono aver motivo di credere che  $R$  sia una convenzione.

Ciò non significa che chi partecipa a una convenzione abbia un modo particolare e infallibile di acquisire le proprie conoscenze. Ma, in un qualche modo ordinario, egli deve *averle acquisite*, in modo da essere uno di coloro tra i quali la convenzione vale. La scoperta della convenzione è l'aspetto principale dell'iniziazione di qualcuno a essa.

Consideriamo le convenzioni di linguaggio, quali che possano essere. Chiunque sia membro di una popolazione P, e sia partecipe delle sue convenzioni di linguaggio, deve sapere quali sono quelle convenzioni. Se in P qualunque regolarità R è di fatto una convenzione di linguaggio, qualunque membro normale<sup>1</sup> di P deve aver motivo di credere che R soddisfi le condizioni definitorie di una convenzione.

Ecco una specie di argomento a difesa della dottrina di Stanley Cavell secondo la quale un parlante nativo non ha bisogno di prove per giustificare ciò che dice intorno a ciò che diremmo. Prendiamo un filosofo che sostiene che noi non chiameremmo volontaria un'azione se non fosse anormale. Egli non ha bisogno di citare occasioni in cui azioni normali non sono state chiamate volontarie, in quanto è un parlante nativo della lingua di cui sta parlando. Perché è giustificato? Non perché, in quanto parlante nativo, possiede qualche modo peculiare e infallibile di acquisire la conoscenza della propria lingua. E non perché la sua conoscenza di ciò che diremmo non sia una conoscenza effettiva, come sembra pensare Cavell quando dice: "il parlante nativo può fare assegnamento su se stesso; altrimenti non ci sarebbe niente da considerare. Infatti la persona che dice quello che diremmo *non* sta soltanto parlando per se stessa."<sup>2</sup>

Semmai, è perché la conoscenza che Cavell ha in mente è la conoscenza che il parlante ha di convenzioni delle quali è partecipe egli stesso. Quando Cavell parla della nostra conoscenza di "ciò che diremmo", io penso che intenda la nostra conoscenza di ciò che *potremmo* dire: che potremmo dire senza violare le nostre convenzioni di linguaggio. Egli non intende la nostra conoscenza di ciò che diremmo per fornire al nostro uditorio le informazioni che desidera; di ciò che diremmo per non essere scortesii o noiosi; di ciò che diremmo tanto per dire.

Una volta che ci siamo resi conto che qualcuno è un parlante nativo della nostra lingua ci siamo già assicurati che è partecipe delle nostre convenzioni. Egli quindi sa che cosa prescrivono quelle convenzioni; sa "ciò che diremmo"

---

<sup>1</sup>Non contando i bambini e gli idioti, che possono conformarsi a R senza aspettarsi la conformità e senza preferire conformarsi subordinatamente alla conformità da parte degli altri.

<sup>2</sup>"Must We Mean What We Say?", *Ordinary Language: Essays in Philosophical Method*, ed. Vere Chappell, Prentice-Hall, Englewood Cliffs, 1964, pp. 75-112.

nel senso in questione. Se a nostra volta gli chiediamo di fornire prove di ciò che dice riguardo a ciò che diremmo, mettiamo in dubbio le sue prerogative di parlante nativo e di partecipe delle convenzioni. Non mettiamo in dubbio qualche prerogativa ulteriore (che egli potrebbe rivendicare) di autorità per quanto riguarda le convenzioni e *anche* la sua prerogativa di soggetto che ne è partecipe. Egli possiede le prove: prove perfettamente ordinarie. Ma se gli chiediamo di esibirle, mettiamo in dubbio la sua appartenenza alla comunità linguistica alla quale egli pretende di appartenere. Non ha senso chiedergli delle prove di ciò che egli dice delle convenzioni e *nello stesso tempo* dare per scontato che egli sia partecipe di quelle convenzioni.

Questa difesa della dottrina di Cavell è tuttavia una difesa assai misera, perché la nostra conoscenza delle nostre convenzioni - quel minimo di conoscenza che ciascuno ha in virtù della sua stessa partecipazione - può essere una specie di conoscenza assai misera:

(1) Può essere una conoscenza puramente potenziale. Dobbiamo disporre di prove in base alle quali possiamo giungere alla conclusione che qualunque nostra convenzione soddisfa le condizioni definitorie di una convenzione, ma può darsi che non abbiamo fatto il ragionamento che ci permette di arrivare a quella conclusione. Se ci viene chiesto se qualcosa è una convenzione, può darsi che diamo un giudizio disinvolto anziché vagliare le nostre prove; quindi può darsi che diamo la risposta sbagliata.

(2) Può essere una conoscenza irrimediabilmente non-verbale. Pensiamo ai rematori della barca di Hume, esempio (3) del capitolo 1.5. Se io sono uno dei rematori e remo con un certo ritmo in virtù di una convenzione tacita e temporanea, ho la prova che abbiamo una convenzione per remare con quel ritmo. Il fatto che nelle ultime battute siamo riusciti a vogare con quel ritmo è una prova in base alla quale io arrivo all'aspettativa che tu continuerai a remare così. Ed è una prova del fatto che tu ti basi su questa stessa prova. Io posso usare tale prova, posso aspettarmi che tu la usi, e così via; ma non posso descriverla. Non posso dire come stiamo remando - per esempio, una vogata ogni 2,3 secondi - ma posso continuare a remare in quel modo; posso dire se tu continui a remare in quel modo; in seguito, potrei probabilmente mostrare a qualcuno che ritmo era; sarei sorpreso se tu cominciassi a remare in modo diverso; e così via. Adesso c'è una descrizione che può identificare il modo in cui remiamo. Impieghiamo  $1,4 \pm 0,05$  secondi per la remata e  $0,9 \pm 0,1$  per il ritorno, esercitando, all'inizio di ogni remata una forza massima di  $70 \pm 10$  libbre, muovendo i remi da  $32^\circ \pm 6^\circ$  in avanti a  $29^\circ \pm 4^\circ$  indietro, e così via, con quanti dettagli vogliamo. Ma, mentre remiamo, un tale genere di descrizione non ci serve a niente. Non possiamo produrla né, se in qualche modo viene prodotta, dire se è vera. Avremmo

bisogno di strumenti, e se anche li avessimo non potremmo continuare a remare come facevamo mentre eseguiamo le misurazioni.

Ci piaccia o no, possediamo molte conoscenze che non possiamo formulare verbalmente. E gran parte delle nostre conoscenze, formulate verbalmente o meno, è fondata su prove che non possiamo sperare di riferire. Le nostre credenze si formano sotto l'influsso di impressioni che ci sono state lasciate da un complesso di esperienze passate, ma è solo occasionalmente che queste impressioni ci permettono di riferire l'esperienza che le ha create. Voi probabilmente credete che la Kamchatka esista. La vostra credenza è giustificata, perché è basata su prove: per la massima parte, la vostra esposizione a vari libri e a eventi che confermano l'attendibilità di tali libri. Provatevi, allora, a dare una dimostrazione convincente dell'esistenza della Kamchatka riferendo parti della vostra esperienza. Non c'è alcun motivo perché la nostra conoscenza delle convenzioni debba essere privilegiata in modo particolare. Come ogni altra conoscenza che abbiamo, può essere tacita, o fondata su prove conosciute tacitamente, o su entrambe le cose.

(3) Può essere una conoscenza limitata a casi particolari, presi uno 'per volta. Una regolarità è convenzionale in virtù di certe aspettative e preferenze generali riguardanti la conformità a essa. Ma non è necessario che queste ultime siano generali in *sensu composito*; è sufficiente la generalità in *sensu diviso*.

La distinzione, risalente a Abelardo, è la seguente. Se io mi aspetto, in *sensu composito*, che ogni guidatore tenga la destra, allora ho una aspettativa di contenuto generale: mi aspetto che ogni guidatore tenga la destra. Non ne segue che, se Jones è un guidatore, io mi aspetti che *egli* tenga la destra, in quanto potrei non capire che è un guidatore. Anzi, potrei addirittura capire che Jones è un guidatore eppure non aspettarmi che egli tenga la destra, perché potrei mancare di trarre le debite conclusioni dalla mia aspettativa generale. Se invece mi aspetto, in *sensu diviso*, che ogni guidatore tenga la destra, allora ho molte aspettative, ciascuna di contenuto *non*-generale. Io mi aspetto *da* Jones, che è un guidatore, che egli tenga la destra. Da Morgan, pure. E così via, per tutti i guidatori che ci sono. Non ho bisogno di sapere che Jones, Morgan e gli altri sono tutti i guidatori che ci sono; potrei pensare falsamente che ci sono altri guidatori che non tengono la destra. La generalità in *sensu composito* e la generalità in *sensu diviso* sono compatibili e spesso coesistono; ma è possibile avere l'una senza l'altra.

La generalità in *sensu diviso* è problematica perché le aspettative e simili si applicano fondamentalmente a stati di cose. Se mi aspetto che ciascun guidatore tenga la destra, mi aspetto uno stato di cose: ciascun guidatore

terrà la destra. Ma se io mi aspetto, *da* ciascun guidatore, che *egli* tenga la destra, quale stato di cose mi aspetto? “*Egli* terrà la destra” non specifica *alcuno* stato di cose finché il pronome non è stato rimpiazzato da qualche specie di descrizione - verbale, raffigurativa o di altro tipo - della persona in questione. Supponiamo che la descrizione “il guidatore della Cadillac color pulce davanti a me” si adatti a *x*. In tal caso posso aspettarmi, *da x*, che *egli* tenga la destra, in quanto ho un’aspettativa che si applica a *x* grazie alla descrizione che gli si riferisce: mi aspetto che il guidatore della Cadillac color pulce davanti a me tenga la destra. In questo caso, c’è uno stato di cose che mi aspetto. Ma non qualunque descrizione di *x* funzionerà. Supponiamo che, a mia insaputa, *x* sia il capo della polizia e anche l’ubriacone della città. Io non mi aspetto, *da x*, che *egli* tenga la destra soltanto perché mi aspetto che il capo della polizia tenga la destra. Né cesso di aspettarmi, *da x*, che *egli* tenga la destra soltanto perché non mi aspetto che l’ubriacone della città tenga la destra. È necessario che la mia aspettativa si applichi a *x* in virtù di una descrizione di qualche specie particolare; ed è difficile dire quali descrizioni andranno bene, e perché.<sup>3</sup>

Consideriamo il caso generale: io mi aspetto che ogni membro di *P* coinvolto con me in un esempio di *S* si conformi a *R*. Abbiamo due quantificazioni universali: una sugli esempi di *S* e una sui membri di *P* coinvolti in un qualunque esempio. È possibile, naturalmente, che la mia aspettativa sia generale in *sensu diviso* riguardo agli esempi di *S*, ma generale in *sensu composito* riguardo agli agenti di un qualunque esempio. Cioè, può darsi che io mi aspetti, da qualunque esempio di *S* in cui sono coinvolto, che tutti in esso si conformino a *R*. (Naturalmente non è possibile che la mia aspettativa sia generale in *sensu composito* riguardo agli esempi e in *sensu diviso* riguardo agli agenti.)

La stessa distinzione tra tipi di generalità si applica a altri atteggiamenti. Prendiamo le nostre preferenze condizionali per la conformità a una convenzione: per esempio, la mia preferenza a conformarmi a *R* in esempi di *S* fra membri di *P*. Se la mia preferenza è generale in *sensu composito* riguardo a esempi di *S*, allora io preferisco lo stato di cose in cui mi conformo a *R* ogniqualvolta sono implicato in un esempio di *S* (fra membri di *P* che si conformano a *R*) allo stato di cose in cui a volte non mi conformo a *R* quando

---

<sup>3</sup>Cfr. Quine, “Quantifiers and Propositional Attitudes”; Richard Montague e Donald Kalish, “That”, *Philosophical Studies*, 10 (1959), pp. 54-61; David Lewis, “Counterpart Theory and Quantified Modal Logic”, *Journal of Philosophy*, 65 (1968), pp. 113-126; David Kaplan, “Quantifying In”, *Synthese*, 19 (1969). [Trad. it. di E. Bencivenga, “Quantificazioni in contesti intermedi, in *Riferimento e modalità*, a cura di L. Linsky, Bompiani, Milano, in corso di stampa. (N.d.T.)]

sono implicato in un esempio di S (fra membri di P che si conformano a R). Ma se la mia preferenza è generale in *sensu diviso*, allora, per qualunque esempio di S (fra membri di P) nel quale sono implicato, preferisco lo stato di cose in cui io e gli altri ci conformiamo a R in quell'esempio allo stato di cose in cui gli altri si conformano a R in quell'esempio e io no. Di nuovo, i due tipi di generalità possono coesistere, e spesso coesistono, ma sono indipendenti.

Quale tipo di generalità riguardo agli esempi di S è necessario nella definizione di convenzione? Io direi: qualunque tipo, purché assicuri la capacità dell'agente di applicare i propri atteggiamenti generali all'esempio in questione. E questa è una generalità limitata in *sensu diviso*. Ogniqualvolta l'agente si trova in un esempio di S fra membri di, P, egli deve aspettarsi che gli altri si conformino a R *in quell'esempio*, preferire di conformarsi a R se essi lo preferiscono *in quell'esempio*, e così via, in modo da aver motivo di conformarsi a R egli stesso. Atteggiamenti generali in *sensu composito* sarebbero un'integrazione conveniente e gradita, e potrebbero servire come fonti di atteggiamenti generali in *sensu diviso*. Ma non sarebbero sufficienti da soli; l'agente dovrebbe essere capace di riconoscere gli esempi di S e di derivare gli atteggiamenti particolari appropriati. Se ciò avvenisse, i suoi atteggiamenti - o almeno la sua propensione ad acquisire atteggiamenti - sarebbero generali in *sensu diviso*.

Possiamo immaginare in che modo una convenzione R riguardante l'azione in S potrebbe valere in una popolazione P di creature incapaci di avere atteggiamenti generali in *sensu composito*. Esse imparano dall'esperienza non già arrivando a credere a certe generalizzazioni, ma acquisendo propensioni a adottare le credenze particolari giuste circa qualunque nuovo caso che sia presentato in modo sufficientemente dettagliato. Esse sono esposte ai precedenti: quelli che noi chiameremmo (ma esse non potrebbero farlo) esempi di S, nei quali venivano conseguiti risultati soddisfacenti per tutti grazie a ciò che noi chiameremmo (ma esse non potrebbero farlo) conformità a R. In seguito, ogniqualvolta si trova davanti a un nuovo esempio di S, anche se non molto simile a un qualsiasi precedente, una di queste creature possiede tutti gli atteggiamenti appropriati relativi a quell'esempio. Essa si aspetta che gli agenti siano reciprocamente impegnati a realizzare quella che noi chiameremmo (ma essa non potrebbe farlo) conformità a R. Considerando due risultati qualunque, la nostra creatura ha una preferenza; e il suo sistema di preferenze tra risultati diversi è tale che c'è un equilibrio di coordinazione in cui tutti gli interessati si conformano a R. Infine, è conoscenza comune tra i membri di P che questi atteggiamenti sono presenti in tutti quelli che prendono parte a *questo particolare* esempio di S. C'è

un stato di cose A tale che, per questo o per qualunque altro esempio di S, per chiunque vi prenda parte, A indica che egli ha, in quell'esempio, gli atteggiamenti appropriati.

Tale creatura ha una convenzione e la conosce in questa misura: dato un qualunque esempio di S, sa come essa stessa e ciascuno dei suoi compagni si comporterebbero riguardo a quell'esempio (cioè, in un qualche modo che noi chiameremmo conformità a R). E sa che essi si comportano così per convenzione; cioè, data una qualunque delle condizioni definitorie di convenzione applicata a un dato agente in una data situazione, essa sa che la condizione è soddisfatta. Ma non può pensare a più di un esempio - quello dato - per volta. La nostra creatura non dispone di alcun concetto generale di esempio di S, di membro di P, o di azione conforme a R.

Supponiamo che noi, che *abbiamo* la capacità di generalizzare, vogliamo utilizzare la conoscenza che questa creatura ha della propria convenzione al fine di dare una descrizione generale della convenzione stessa. Dovremo procedere per prova ed errore, escogitando ipotesi e verificandole su di essa in un esempio (ben scelto) dopo l'altro, fino a quando pensiamo di essere in grado di predire la sua risposta a qualunque esempio futuro. Anche a noi, cui sarebbe possibile conoscere le nostre convenzioni generalmente in *sensu composito*, potrebbe accadere di conoscerle generalmente soltanto in *sensu diviso*. Se volessimo conoscerle generalmente anche in *sensu composito*, dovremmo far ricorso allo stesso tipo di prova ed errore, con noi stessi come soggetti. I nostri dati circa gli esempi delle nostre convenzioni sarebbero attendibili. Ma le nostre ipotesi generali volte alla sistematizzazione di quei dati sarebbero normali ipotesi sperimentali, prive di uno status privilegiato.

### 2.3 Alternative alle convenzioni

Una delle mie condizioni definitorie del carattere convenzionale di una regolarità R relativa alla scelta di azione da parte di agenti in una situazione S era:

In qualunque esempio di S fra membri di P, tutti preferiscono conformarsi a R a condizione che lo facciano gli altri, in quanto S è un problema di coordinazione e la conformità uniforme a R è un equilibrio di coordinazione in S.

Nella discussione dell'esempio (7) del capitolo 1.5 abbiamo trovato questa condizione insoddisfacente ogniqualvolta avevamo non solo coordinazione di azioni in un esempio qualunque di S, ma coordinazione tra azioni in esempi

assai vicini di  $S$  all'interno di una qualche attività continua. E abbiamo visto che il rimedio non consisteva nel considerare come problemi di coordinazione periodi di attività più lunghi, in quanto periodi più lunghi non sono problemi di coordinazione. Adesso formulerò delle nuove condizioni che differiscono dalla precedente solo per il fatto che non richiedono che la nostra attività sia segmentata in problemi di coordinazione indipendenti. Le nostre nuove condizioni non implicheranno che  $S$  sia un problema autonomo di decisione interdipendente, nel quale ciascun agente fa una scelta d'azione e il risultato che ciascuno consegue dipende dalle azioni di tutti; ma implicheranno che, *se*  $S$  è un tale problema, *allora*  $S$  è un problema di coordinazione e la conformità uniforme a  $R$  è un equilibrio di coordinazione in  $S$ . Il caso particolare di una sequenza di problemi di coordinazione sarà trattato come sopra; ma ci accorgeremo che, nello stesso tempo, ci siamo occupati anche degli altri casi.

In primo luogo, richiediamo che ciascun agente implicato in un esempio di  $S$  preferisca conformarsi a  $R$  subordinatamente alla conformità da parte degli altri agenti che sono implicati con lui in  $S$ . Egli preferisce la conformità uniforme a  $R$  a qualunque combinazione di azioni in cui gli altri si conformano ed egli no. Se  $S$  è un problema autonomo di decisione interdipendente, il primo requisito fa della conformità uniforme a  $R$  un equilibrio. Altrimenti, la conformità uniforme non è un equilibrio ma qualcosa che gli assomiglia molto.

In secondo luogo, richiediamo che tutti gli agenti implicati abbiano pressappoco le stesse preferenze per quanto riguarda le combinazioni delle loro azioni, in modo che  $S$  sia una situazione in cui predomina la coincidenza di interessi. In particolare, richiediamo che tutti condividano la preferenza condizionale di ciascuno per la propria conformità a  $R$ . Cioè, come io preferisco conformarmi se tu e gli altri vi conformate, così anche tu preferisci che io mi conformi se tu e gli altri vi conformate. Se unifichiamo questa condizione con la precedente, otteniamo: ciascuno preferisce che tutti si conformino a  $R$ , a condizione che almeno tutti tranne uno si conformino a  $R$ , sia che quell'uno sia egli stesso sia che si tratti di qualcun altro. Se  $S$  è un problema autonomo di decisione interdipendente, questo secondo requisito fa della conformità uniforme a  $R$  un equilibrio di coordinazione proprio.

Infine, richiediamo che ci sia una seconda regolarità possibile  $R'$  (riguardante la scelta di azioni da parte degli agenti in  $S$ ) che soddisfa le stesse condizioni che imponiamo a  $R$ . Chiamiamo  $R'$  alternativa a  $R$ . È sufficiente richiedere che  $R'$  soddisfi la prima e la seconda condizione imposta a  $R$ . La terza è automatica: se  $R$  ha come *alternativa*  $R'$ , allora  $R'$  ha come alternativa  $R$ . Se  $S$  è un problema autonomo di decisione interdipendente,

quest'ultimo requisito fa della conformità uniforme a  $R'$  un secondo equilibrio di coordinazione proprio. Perciò esso garantisce che  $S$  soddisfa l'ultima condizione che definisce un problema di coordinazione: il possesso di due o più equilibri di coordinazione propri.

Ricordiamo la discussione del capitolo I.2 sulla banalità di qualunque situazione con un equilibrio di coordinazione unico e interessi prevalentemente coincidenti. Adesso ci è più agevole descrivere tale banalità: la conoscenza comune della razionalità è tutto ciò di cui un agente ha bisogno per aver motivo di svolgere la propria parte in un equilibrio di coordinazione. Egli non ha bisogno di fare appello a precedenti o a qualunque altra fonte di ulteriori aspettative reciproche.

Finora abbiamo sottratto la convenzione a questa banalità richiedendo che  $S$  sia un problema di coordinazione, avente quindi, per definizione, più di un equilibrio di coordinazione proprio. Adesso che non richiediamo più che  $S$  sia un problema di coordinazione, la nostra richiesta di un'alternativa si ispira alla stessa strategia. Di fatto, quando  $S$  è un problema autonomo di decisione interdipendente, non abbiamo introdotto alcun cambiamento.

La nostra nuova condizione serve a mettere in evidenza una proprietà delle convenzioni che prima non era stata sottolineata: non esiste qualcosa come l'unica convenzione possibile. Se  $R$  è la nostra effettiva convenzione,  $R$  deve avere l'alternativa  $R'$ , e  $R'$  deve essere tale che avrebbe potuto essere la nostra convenzione invece di  $R$ , se soltanto la gente avesse cominciato a conformarsi a  $R'$  e ad aspettarsi reciprocamente di conformarsi a  $R'$ . È per questo che è ridondante parlare di una convenzione arbitraria. Qualunque convenzione è arbitraria perché c'è una regolarità alternativa che avrebbe potuto essere la nostra convenzione. Una convenzione che, per così dire, *non* sia arbitraria è una regolarità per mezzo della quale raggiungiamo equilibri di coordinazione unici. Dato che non è arbitraria, non deve neppure essere convenzionale. Noi ci conformeremmo ad essa semplicemente perché è la cosa migliore da fare. Indipendentemente da quanto abbiamo fatto nel passato, una mancata conformità alla "convenzione non arbitraria" potrebbe essere soltanto un errore strategico (o la compensazione di un previsto errore strategico di qualcun altro, o la compensazione di una prevista compensazione da parte di qualcun altro ecc.).

Quando cerchiamo di formulare in modo più preciso il requisito dell'alternativa, sorge un problema. Si suppone che  $R$  e  $R'$  siano diverse, e quindi che un'azione conforme a  $R$  (da parte di un agente in  $S$ ) non sia conforme anche a  $R'$ , e viceversa. Ma sono sempre diverse, o sono diverse solo qualche volta? Dopo tutto non è necessario che gli esempi di  $S$  siano esattamente uguali. Devono essere semplicemente analoghi, cadere sotto una qualche de-

scrizione comune che sia abbastanza naturale da permettere la conoscenza comune di una propensione a compiere estrapolazioni da alcuni esempi di S a altri. Così, un'azione conforme a R potrebbe essere conforme anche a R' per alcuni agenti in alcuni esempi di S, anche se non per tutti.

	C1	C2
R1	1	0
R2	0,5	0,2

Figura 27

	C3	C4
R3	1	0
R4	0,2	0,5

Figura 28

Non è sufficiente richiedere che l'alternativa R' differisca da R solo nella misura in cui è incompatibile con R per alcuni agenti, o anche per tutti gli agenti, in alcuni esempi possibili. Supponiamo che S occorra in una versione frequente, illustrata nella Figura 27, e in una versione rara, illustrata nella Figura 28. (Trascuriamo ogni ulteriore differenza tra esempi di una versione.) S è banale in una versione ma non nell'altra. Sia R la regolarità consistente nel fare R1 o C1 (nella versione frequente) oppure R3 o C3 (nella versione rara). Ritengo che R non andrebbe definita come una convenzione, in quanto è banale nella maggior parte degli esempi di S. Ma avrebbe le caratteristiche di una convenzione se considerassimo R' come sua alternativa,

dove  $R'$  è la regolarità consistente nel fare  $R1$  o  $C1$  oppure  $R4$  o  $C4$ .  $R'$  è una regolarità desiderabile che è incompatibile con  $R$  in alcuni esempi di  $S$ .

Sarebbe meglio richiedere un'alternativa  $R'$  che sia uniformemente incompatibile con  $R$ , incompatibile per tutti gli agenti in tutti gli esempi di  $S$ . Ora,  $R$ , nell'esempio precedente, non va bene. La sua unica alternativa uniformemente incompatibile sarebbe  $R''$ , la regolarità consistente nel fare  $R2$  o  $C2$  oppure  $R4$  o  $C4$ . Ma  $R''$  non è un'alternativa a  $R$ , in quanto  $R''$  non soddisfa di solito il requisito della preferenza condizionale per la conformità. Negli esempi della versione frequente di  $S$  nessuno desidera conformarsi a  $R''$  anche se il partner lo fa.

Se ogni esempio di  $S$  è un problema di coordinazione, e se la conformità uniforme a  $R$  è sempre un equilibrio di coordinazione proprio, allora possiamo trovare un altro equilibrio di coordinazione proprio in ogni esempio di  $S$ . Perciò la regolarità  $R'$ , in virtù della quale uno svolge la propria parte in un secondo equilibrio di coordinazione proprio prescelto in ogni esempio di  $S$ , è un'alternativa a  $R$ , e  $R$  e  $R'$  sono uniformemente incompatibili.

Se lo preferiamo, tuttavia, non dobbiamo necessariamente richiedere un'alternativa uniformemente incompatibile con  $R$ . Una versione (apparentemente) più debole consisterebbe nel richiedere soltanto che per ogni esempio di  $S$  ci sia un'opportuna regolarità  $R'$  che sia incompatibile con  $R$  (per tutti gli agenti implicati) in *quell'*esempio. Alternative a  $R$  parzialmente incompatibili sono sufficienti se è sufficiente il loro numero. Le due versioni non sono in realtà diverse. La versione forte implica direttamente la versione debole; e la versione debole implica indirettamente la versione forte, in quanto possiamo sempre ottenere un'alternativa uniformemente incompatibile mettendo insieme i pezzi di alternative parzialmente incompatibili.

Potremmo quindi sostituire la nostra condizione iniziale con due nuove. Questa:

In qualunque esempio di  $S$  tra membri di  $P$ , tutti hanno pressappoco le stesse preferenze riguardo a tutte le combinazioni possibili di azioni; assieme a questa (versione forte): Nel comportamento dei membri di  $P$  in  $S$  c'è qualche regolarità possibile  $R'$  tale che nessuno, in qualunque esempio di  $S$  tra membri di  $P$ , potrebbe conformarsi sia a  $R'$  che a  $R$ , e tale che, in qualunque esempio di  $S$  fra membri di  $P$ , tutti preferirebbero che tutti si conformassero a  $R'$ , a condizione che almeno tutti tranne uno si conformassero a  $R'$ ; o a questa (versione debole): In qualunque esempio di  $S$  fra membri di  $P$  c'è, nel comportamento dei membri di  $P$  in  $S$ , qualche regolarità possibile  $R'$  tale che nessuno

in quell'esempio di S potrebbe conformarsi sia a R' che a R, e tale che tutti preferirebbero che tutti si conformassero a R', a condizione che almeno tutti tranne uno si conformassero a R'.

Non vedo possibilità di scegliere l'una piuttosto che l'altra versione, e scelgo quella forte senza un motivo particolare.

Quando S è un problema autonomo di decisione interdipendente, le nostre nuove condizioni concordano con quella iniziale. Ma queste non richiedono che S sia autonomo. Altrimenti (come nel mio esempio della definizione dei prezzi, con S interpretato come un periodo di attività commerciale abbastanza lungo da includere più decisioni relative ai prezzi) le nuove condizioni sono un'estensione naturale della condizione iniziale.

Sia R una convenzione relativa al comportamento in un problema di coordinazione S; e sia R' un'altra regolarità possibile, parzialmente o uniformemente incompatibile con R, che risolverebbe S. Ma supponiamo che l'equilibrio di coordinazione che raggiungeremmo conformandoci a R' sia molto peggiore di quello che raggiungiamo conformandoci a R: tanto peggiore, in effetti, che è preferito solo di poco ad alcuni degli esiti che non sono equilibri di coordinazione. Vorremo davvero chiamare R' una convenzione alternativa possibile? E vorremo dire che R' contribuisce all'arbitrarietà e alla convenzionalità di R? Forse no. Fortunatamente la nostra definizione, così com'è, è adatta a escludere questa R' come alternativa a R. Può essere vero che:

In qualunque esempio di S fra membri di P, tutti preferirebbero che tutti si conformassero a R', a condizione che almeno tutti tranne uno si conformassero a R'.

Ma ciò può non essere vero come elemento di conoscenza comune. Le nostre preferenze condizionali deboli per la conformità a R' possono benissimo non essere indicate da alcuno stato di cose A che noi tutti crediamo valere e che ci indica di valere. Ma naturalmente richiediamo ancora il soddisfacimento delle condizioni per essere conoscenza comune in P. E a ragione: se la nostra preferenza condizionale per la conformità a R' esistesse, ma non come conoscenza comune, R' non avrebbe potuto consolidarsi allo stesso modo di una convenzione, per cui non è vero che R' avrebbe potuto essere la nostra convenzione al posto di R.

Consideriamo una convenzione che stabilisce un qualche luogo di incontro. Sue alternative sarebbero le regolarità possibili in virtù delle quali ci incontreremmo in altri posti. Alcuni luoghi sono meglio di altri. Alcuni vanno così male che rinunceremmo a incontrarci piuttosto che andarci.

Considerando luoghi sempre peggiori, giungiamo al punto in cui la preferenza condizionale per la conformità viene meno: alcuni di noi non vorrebbero andare in un posto anche se gli altri fossero là. La conoscenza comune della preferenza condizionale viene meno prima: ci sono luoghi che vanno abbastanza bene perché ciascuno desideri recarvisi se vi si recano gli altri, ma che non vanno abbastanza bene perché possiamo fare assegnamento sulla nostra reciproca disposizione a recarci là se vi si recano gli altri, fare assegnamento sul fatto che facciamo reciprocamente assegnamento sulla nostra reciproca disposizione a recarci là se vi si recano gli altri, e così via. Questi luoghi non ci forniscono alternative alla nostra convenzione, per cui non contribuiscono alla convenzionalità del nostro luogo d'incontro.

Oppure consideriamo le convenzioni della nostra lingua. Le loro alternative sono le convenzioni di altre lingue possibili. Ma che dire I. una lingua ipotetica - dobbiamo chiamarla cifrario? - così grossolana che anche dopo una qualunque quantità di esercizi ci vorrebbero dei minuti di calcoli con carta e matita per costruire o interpretare i suoi enunciati più facili? *Potrebbe* darsi che noi tutti trovassimo perfino quella lingua meglio che niente, degna di essere appresa per essere usata con chi la usasse. Ma se non fosse conoscenza comune che abbiamo tale opinione di essa, questa lingua non figurerebbe tra le alternative che fanno della nostra lingua effettiva una convenzione.

Se le uniche alternative a R fossero di questa specie difettosa, R non sarebbe una convenzione. Né sarebbe stata una convenzione in base alla condizione iniziale che richiedeva che S fosse un problema di coordinazione, perché neanche quella sarebbe stata conoscenza comune. E non dovrebbe essere chiamata convenzione. Se le alternative a R sono un aspetto così irrilevante della situazione, R sembra quasi altrettanto banale che se quelle alternative non ci fossero affatto.

Può R' essere un'alternativa a R se l'idea di comportarsi in conformità a R' non è mai venuta in mente a nessuno in P? Il principio è lo stesso: la mancanza di familiarità esclude R.' se e solo se interferisce con la conoscenza comune della preferenza condizionale per la conformità a R'. Può escluderla e non escluderla. Dato che le aspettative e le preferenze menzionate nella definizione di convenzione devono soltanto essere generali in *sensu diviso*, non importa se non abbiamo alcun concetto generale di azione conforme a R'. In effetti, nell'ultimo paragrafo ho sostenuto che non sorgerebbe alcun problema neppure se non avessimo alcun concetto generale di azione conforme alla nostra convenzione effettiva. La mancanza di familiarità acquisterebbe invece importanza se ci inducesse a non apprezzare i vantaggi di una qualche azione conforme a R' quando ci trovassimo di fronte a un particolare

esempio di S in cui gli altri si conformassero a R'; o se provocasse il venir meno della conoscenza comune che quel vantaggio *verrebbe* apprezzato.

Di nuovo, non c'è alcun contrasto con la condizione iniziale, quando questa si applica. Se le uniche alternative a R venissero escluse per la loro mancanza di familiarità, non sarebbe conoscenza comune il fatto che S era un problema di coordinazione. R non sarebbe quindi una convenzione in base a nessuna delle due condizioni, anche se potrebbe diventarlo allorquando i membri di P diventassero consapevoli della possibilità di comportarsi conformemente a R'.

Ciò che non è convenzionale tra persone dalla mente limitata e rigida, che non saprebbero cosa fare se gli altri si comportassero diversamente, può essere convenzionale tra persone più duttili. Ciò che non è convenzionale può diventarlo quando giunge notizia di altri che si comportano diversamente; o quando qualcuno inventa un nuovo modo di comportarsi, sia pure un nuovo modo che nessuno adotta. Quando i bambini e gli idioti si conformano alle nostre convenzioni, può darsi che non partecipino a esse in quanto convenzioni, perché può darsi che manchino di preferenze condizionali per la conformità a una alternativa; oppure può darsi che abbiano le preferenze appropriate, ma non come conoscenza comune. Non trovo che questi corollari della nostra analisi della convenzione siano né graditi né sgraditi. L'analisi risponde a problemi finora lasciati irrisolti.

Se sembra ragionevole escludere alternative che sono troppo insoddisfacenti o troppo poco familiari, in quanto non contribuiscono all'arbitrarietà di una convenzione, abbiamo una nuova ragione di richiedere la conoscenza comune. Poiché è per mezzo del requisito della conoscenza comune che possiamo escluderle con un procedimento non *ad hoc*.

## 2.4 Gradi di convenzione

Abbiamo limitato la nostra attenzione a casi perfetti di convenzione, ai quali la nostra definizione si applica senza eccezioni. Ma non possiamo sperare di trovare nella realtà molti esemplari perfetti. È tempo di essere meno severi, di lasciare aperta la possibilità di convenzioni che soddisfano la presente definizione solo per la massima parte o con grande probabilità. Riscriviamo la definizione con tutte le correzioni introdotte finora:

Una regolarità R nel comportamento dei membri di una popolazione P, quando essi sono agenti in una situazione ricorrente S, è una *convenzione* se e solo se è vero che, e in P è conoscenza comune che, in qualunque esempio di S tra membri di P,

- (1) tutti si conformano a R;
- (2) tutti si aspettano che tutti gli altri si conformino a R;
- (3) tutti hanno pressappoco le stesse preferenze riguardo a tutte le combinazioni possibili di azioni;
- (4) tutti preferiscono che tutti si conformino a R, a condizione che almeno tutti tranne uno si conformino a R;
- (5) tutti preferirebbero che tutti si conformassero a R' a condizione che almeno tutti tranne uno si conformassero a R',

dove R' è una qualche regolarità possibile del comportamento di membri di P in S tale che nessuno, in qualunque esempio di S fra membri di P, potrebbe conformarsi sia a R' che a R.

Possiamo enumerare diverse quantificazioni universali esplicite e implicite; vogliamo trovare una maniera ragionevole di ridurle, in parte o tutte, a quantificazioni quasi-universali.

Il requisito della conoscenza comune contiene quantificazioni universali su P (cfr. la definizione di conoscenza comune). Non è necessario che ammettiamo delle eccezioni a esse; si potrebbe piuttosto escludere da P chiunque possa essere chiamato un'eccezione. Ne segue, comunque, che la maggior parte delle nostre specificazioni relative a una popolazione nella quale vale una convenzione saranno corrette soltanto in misura approssimativa.

Non c'è alcun male nel permettere alcuni esempi anormali di S che violano alcune o tutte le clausole (1)-(5). Sostituiamo quindi "in qualunque esempio di S fra membri di P" con "in quasi tutti gli esempi di S tra membri di P". Se desideriamo una precisione maggiore, possiamo sostituirlo con "in una frazione di almeno  $d_0$  di tutti gli esempi di S fra membri di P", dove il valore di  $d_0$  è fissato leggermente al di sotto di uno.

Né c'è alcun male nel permettere che alcuni, o anche la maggior parte, degli esempi normali di S contengano alcuni agenti anormali che, in alcune o in tutte le clausole (1)-(5), possono essere eccezioni alle quantificazioni universali iniziali. Sostituiamo quindi ogni "tutti" posto all'inizio con "quasi tutti" o con "tutti in una frazione di almeno  $d_i$  di tutti gli agenti", dove il valore di ogni  $d_i$  è fissato leggermente al di sotto di uno. (Abbiamo  $d_1$  per la clausola (1),  $d_2$  per la clausola (2),  $d_3$  per la clausola (3) e  $d_4$  per le clausole (4) e (5): lo stesso per entrambe perché sono concepite come parallele.)

Se permettiamo che ci siano alcuni agenti che non si conformeranno, dovremmo permettere che gli altri agenti lo sappiano; così, nella clausola (2), “tutti gli altri” andrebbe sostituito con “quasi tutti gli altri” o con “tutti gli altri in una frazione di almeno  $d_1$  di tutti gli agenti”. E se permettiamo che gli agenti non si aspettino una conformità perfetta, non dobbiamo far sì che le loro preferenze per la conformità siano subordinate a una conformità altrimenti perfetta; in caso contrario non ci sarebbe garantito che essi preferiscano la conformità nella maggior parte dei casi. Le loro preferenze dovrebbero essere tali che, se un numero sufficiente si conforma, allora quanti più si conformano tanto meglio è. (Una cosa che non tolleriamo è quindi una convenzione alla quale la maggior parte delle persone vuole che ci siano eccezioni, per quanto poche possano essere le eccezioni desiderate.) La clausola (4) andrebbe quindi corretta di nuovo nel modo seguente: “preferiscono che chiunque altro si conformi a R, a condizione che quasi ogni altro si conformi a R”, oppure “preferiscono che chiunque altro si conformi a R, a condizione che una frazione di almeno  $d_1$  di tutti gli agenti si conformi a R”. Sebbene questa correzione renda la clausola (4) più rigida anziché meno rigida, essa è inevitabile, dato l’indebolimento delle clausole (1) e (2) da noi effettuato. La clausola (5) andrebbe corretta nella stessa maniera, in modo da esser mantenuta parallela a (4).

Possiamo tollerare alcune eccezioni anche alla richiesta incompatibilità tra R e l’alternativa R’: eccezioni per la maggior parte degli agenti in alcuni esempi di S, per alcuni agenti nella maggior parte degli esempi di S, o ambedue. Sostituiamo la clausola di incompatibilità con “tale che quasi nessuno, in quasi ogni esempio di S fra membri di P, potrebbe conformarsi sia a R’ che a R”, oppure con “tale che, per una frazione di almeno  $d_5$  di tutte le coppie formate da un esempio di S fra membri di P e un agente in esso implicato, l’agente non potrebbe conformarsi sia a R’ che a R”, dove il valore di  $d_5$  è fissato leggermente al di sotto di uno.

La nostra definizione finale è quindi:

Una regolarità R nel comportamento dei membri di una popolazione P, quando essi sono agenti in una situazione ricorrente S, è una *convenzione* se e solo se è vero che, e in P è conoscenza comune che, in qualunque esempio di S fra membri di P,

- (1) quasi tutti si conformano a R;
- (2) quasi tutti si aspettano che quasi tutti gli altri si conformino a R;

- (3) quasi tutti hanno pressappoco le stesse preferenze riguardo a tutte le combinazioni possibili di azioni;
- (4) quasi tutti preferiscono che chiunque altro si conformi a R, a condizione che quasi tutti si conformino a R;
- (5) quasi tutti preferirebbero che chiunque altro si conformasse a R', a condizione che quasi tutti si conformassero a R',

dove R' è una qualche regolarità possibile del comportamento di membri di P in S tale che quasi nessuno, in quasi ogni esempio di S fra membri di P, potrebbe conformarsi sia a R' che a R.

Se qualcuno lamenta l'imprecisione della nostra definizione finale di convenzione, è libero di far uso della seguente definizione quantitativa.

Una regolarità R nel comportamento dei membri di una popolazione P, quando questi sono agenti in una situazione ricorrente S, è una *convenzione* perlomeno di gradi  $d_0, d_1, d_2, d_3, d_4, d_5$  se e solo se, in una frazione di almeno  $d_0$  di tutti gli esempi di S fra membri di P, è vero, e è conoscenza comune in P, che

- (1) tutti in una frazione di almeno  $d_1$  di tutti gli agenti si conformano a R;
- (2) tutti in una frazione di almeno  $d_2$  di tutti gli agenti si aspettano che tutti gli altri in una frazione di almeno  $d_1$  di tutti gli agenti si conformino a R;
- (3) tutti in una frazione di almeno  $d_3$  di tutti gli agenti hanno pressappoco le stesse preferenze riguardo a tutte le combinazioni possibili di azioni;
- (4) tutti in una frazione di almeno  $d_4$  di tutti gli agenti preferiscono che chiunque altro si conformi a R, a condizione che una frazione di almeno  $d_1$  di tutti gli agenti si conformi a R;
- (5) tutti in una frazione di almeno  $d_4$  di tutti gli agenti preferirebbero che chiunque altro si conformasse a R', a condizione che una frazione di almeno  $d_1$  di tutti gli agenti si conformasse a R',

dove  $R'$  è una qualche regolarità possibile del comportamento di membri di  $P$  in  $S$  tale che, per una frazione di almeno  $d_5$  di tutte le coppie formate da un esempio di  $S$  fra membri di  $P$  e un agente in esso implicato, l'agente non potrebbe conformarsi sia a  $R'$  che a  $R$ .

Costui può andare avanti a definire la convenzione come qualunque regolarità che sia una convenzione di almeno certi gradi fissati, gradi che può scegliere a piacere.

Definiamo il grado di convenzionalità di una regolarità  $R$  come l'insieme delle sestuple  $\langle d_i \rangle$  tali che  $R$  è una convenzione almeno di gradi  $d_0, d_1, d_2, d_3, d_4, d_5$ . Possiamo confrontare le regolarità in base ai loro gradi di convenzionalità:  $R1$  è *più convenzionale* di  $R2$  se e solo se il grado di convenzionalità di  $R2$  è un sottoinsieme del grado di convenzionalità di  $R1$ . Sarebbe interessante trovare un unico numero che misuri il grado di convenzionalità di una regolarità; ma tutte le procedure che conosco per ottenere questo risultato sembrano molto artificiali. Se  $R$  è una convenzione in base alla definizione stretta posta all'inizio di questo paragrafo, allora  $R$  è una convenzione perlomeno di gradi 1, 1, 1, 1, 1, e nessun'altra regolarità può essere più convenzionale.

## 2.5 Conseguenze delle convenzioni

Supponiamo che  $R$  sia una regolarità convenzionale; e supponiamo che  $R^*$  sia una qualche conseguenza logica di  $R$ . È  $R^*$ , per questo motivo, una convenzione a pieno diritto?

Ci sono conseguenze banali delle convenzioni, e di queste non ci interessiamo. Sia  $R$  la nostra convenzione di guidare a destra; una conseguenza logica di  $R$  è che guidiamo sulla superficie delle strade, non a dieci piedi di altezza o dieci piedi sottoterra. Ancora più banalmente, una tautologia, che è conseguenza di qualunque cosa, è conseguenza di qualunque convenzione. Quelle che vogliamo prendere in considerazione sono le conseguenze delle convenzioni che dipendono dalla convenzione. La nostra conseguenza  $R^*$  *dipende* da  $R$  solo se c'è una regolarità  $R'$  che è un'alternativa a  $R$  (nel senso del paragrafo 3) e *non- $R^*$*  è una conseguenza logica di  $R'$ .

Se è così,  $R^*$  può essere una convenzione. Supponiamo che tu e io vogliamo incontrarci ogni settimana; e supponiamo di trascorrere, alternatamente, una settimana nella città  $C1$  e una settimana nella città  $C2$ . La città  $C1$  ha tre luoghi di incontro accettabili:  $L11, L12$  e  $L13$ . Anche la città  $C2$  ha tre luoghi di incontro accettabili, analoghi ciascuno al luogo di incontro di

CI con lo stesso numero: L21, L22 e L23. La nostra convenzione R è la seguente: nelle settimane che passiamo in CI ci rechiamo in L11, nelle settimane che passiamo in C2 ci rechiamo in L21. Una conseguenza R\* di R è questa: nelle settimane che passiamo in C1 ci rechiamo in L11. Si tratta di una conseguenza dipendente, in quanto *non-R\** sarebbe una conseguenza della maggior parte delle alternative a R. R\* è certamente una convenzione. Nelle situazioni alle quali si applica R\* - le settimane che passiamo in C1 - è conoscenza comune, tra noi, che ci conformiamo a R\*, che ci aspettiamo reciprocamente di conformarci a R\*, e che la conformità uniforme a R\* sia un equilibrio di coordinazione in un problema di coordinazione. In generale, la particolarizzazione di una convenzione è una convenzione. Forse la conseguenza di una convenzione è a pieno diritto una convenzione solo se è una particolarizzazione della convenzione iniziale.

	C1	C2	C3
R1	1	0	0,6
R2	0	1	0,6
R3	0,6	0,6	1

Figura 29

Consideriamo ora una conseguenza dipendente di una convenzione che non è essa stessa una convenzione. Supponiamo che ci sia soltanto una città con tre luoghi di incontro accettabili: L1, L2 e L3. Supponiamo di volerci incontrare ma, nel caso che non ci incontriamo, è desiderabile che uno di noi vada in L3. Supponiamo che la nostra matrice delle vincite sia come nella Figura 29, e supponiamo che la nostra convenzione R sia di andare in L1. Sia R\* la regolarità consistente nel recarsi o in L1 o in L2. R\* è una conseguenza di R; ed è una conseguenza dipendente, in quanto non R\* sarebbe una conseguenza della regolarità consistente nel recarsi in L3, che è un'alternativa a R. Noi ci conformiamo a R\* e ci aspettiamo reciprocamente di farlo, e entrambi questi fatti sono, tra noi, conoscenza comune. Ma R\*

non è una convenzione perché, secondo quanto sostengo, non si dà il caso che ciascuno di noi preferisca conformarsi a  $R^*$  a condizione che gli altri si conformino a  $R^*$ . Ammesso soltanto che tu ti conformerai a  $R^*$ , senza alcuna indicazione se lo farai recandoti in L1 o recandoti in L2, io preferisco violare  $R^*$  recandomi in L3. Lo stesso vale per te rispetto a me.

Tuttavia, la questione non è completamente chiara. Osserviamo che quella che chiamiamo preferenza subordinata a un qualche stato di cose A è quasi sempre subordinata anche a qualche stato di cose di sfondo B, che noi consideriamo come una parte fissa del contesto. Dire che io preferisco guidare a destra se lo fanno gli altri equivale in realtà a dire che io preferisco guidare a destra se lo fanno gli altri e se continuano a valere diversi fatti comuni relativi alle cause e agli effetti delle collisioni. Ora, è un fatto, e tra noi è conoscenza comune, che se uno di noi si conforma a  $R^*$ , lo farà conformandosi a R; in altre parole, è un fatto e una conoscenza comune che non andremo in L2. Se questo fatto appartenesse al contesto fisso, allora ciascuno di noi *preferirebbe* conformarsi a  $R^*$ , subordinatamente al conformarsi a  $R^*$  da parte dell'altro e al contesto. Sono sicuro che è sbagliato inserire nel contesto fisso questo fatto, e cioè che se uno di noi si conforma a  $R^*$  lo farà conformandosi a R. Ma non dispongo di alcuna teoria per spiegare perché ciò sia sbagliato. Grosso modo, la ragione è questa: nel considerare le preferenze per azioni subordinate ad azioni, il contesto andrebbe mantenuto neutrale riguardo al problema di sapere se le azioni della specie generale in esame sono compiute o meno.

## Capitolo 3

# Confronto tra convenzione e altre nozioni

### 3.1 Accordo

Nel capitolo 1.3 abbiamo visto come agenti implicati in un singolo e irripetuto problema di coordinazione potrebbero riuscire a raggiungere la coordinazione stipulando l'esplicito accordo di svolgere ciascuno la propria parte in un certo equilibrio di coordinazione. Ma abbiamo anche detto che l'accordo non era l'unico strumento di coordinazione.

Analogamente, il nostro esplicito accordo di conformarci a qualche opportuna regolarità  $R$  è un mezzo - un buon mezzo, ma non l'unico - per stabilire una volta per tutte di raggiungere la coordinazione ogniqualvolta ci si presenta un certo problema di coordinazione ricorrente  $S$ . Accordandoci esplicitamente di conformarci a  $R$ , ciascuno di noi manifesta la sua propensione (preesistente o appena sorta) a conformarsi a  $R$ ; e fa questo in circostanze in cui tra noi è conoscenza comune che noi tutti lo stiamo tenendo d'occhio. Di conseguenza, diviene conoscenza comune che è probabile che noi tutti ci conformiamo a  $R$ . Siccome è anche conoscenza comune che ciascuno preferisce conformarsi a  $R$  a condizione che si conformino anche gli altri implicati con lui in  $S$ ,<sup>1</sup> ciascuno ha tanti più motivi e tanta maggiore propensione a conformarsi: e anche questa è conoscenza comune. La conoscenza comune di una propensione a conformarsi produce azioni conformi quando si presenta il successivo  $S$ ; l'azione conforme rinnova la

---

<sup>1</sup>Se non lo era prima, diventa conoscenza comune durante la discussione che precede il nostro accordo, quando ciascuno di noi manifesta la propria preferenza condizionale per il conformarsi.

nostra conoscenza comune di una propensione a conformarsi. È in atto una convenzione.

La nostra convenzione è il prodotto del nostro accordo, e tali sono sempre - in un certo senso - tutte le nostre azioni conformi. Ma dire che ci comportiamo come ci comportiamo perché una volta ci siamo accordati sarebbe assai fuorviante. Ciò suggerisce che il nostro accordo continui a influenzare direttamente le nostre azioni, esattamente come faceva all'inizio; in realtà il suo effetto principale viene trasmesso attraverso una crescente catena causale di aspettative, azioni, aspettative, azioni e così via. L'influenza diretta svanisce con il passare dei giorni, degli anni, o della vita. Ci dimentichiamo del nostro accordo. Cessiamo di sentirci vincolati dall'antica promessa (se il nostro accordo *era* uno scambio di promesse; come abbiamo visto nel caso di un singolo problema di coordinazione, di solito uno scambio di manifestazioni della nostra intenzione attuale è abbastanza adatto). Noi lasciamo la popolazione, e siamo sostituiti da eredi che non hanno partecipato all'accordo. Ma l'influenza indiretta dell'accordo si rinnova continuamente, e con il passare del tempo diventa predominante. A questo punto una convenzione sorta per accordo non differisce più da una nata diversamente: non conserva più traccia della propria origine.

Di fatto, una convenzione sorta per accordo può non diventare una convenzione, in base alla mia definizione, fino a che l'influenza diretta dell'accordo non ha avuto tempo di svanire. Ciò dipende dalla natura dell'accordo. Supponiamo di aver prestato tutti il solenne e pubblico giuramento di conformarci a R, qualunque cosa potesse accadere. Allora, per un certo tempo, noi tutti potremmo preferire incondizionatamente di conformarci a R, decisi ciascuno a mantenere il proprio giuramento anche se gli altri dovessero infrangerlo per conformarsi a qualche regolarità alternativa R'. Anche se le nostre preferenze per la conformità a R fossero di fatto subordinate alla conformità da parte degli altri, ciò potrebbe non essere conoscenza comune. Infatti il peso del giuramento infranto potrebbe creare incertezza, ci si potrebbe aspettare che crei incertezza, e così via. Abbiamo una convenzione solo dopo che il potere della promessa è svanito al punto che è vero e è conoscenza comune che, se gli altri lo facessero, ciascuno si conformerebbe, anziché a R, a qualche regolarità alternativa R'.

Se, invece, ci fossimo accordati scambiandoci promesse condizionali che ci vincolavano a conformarci a R solo se lo facevano gli altri, oppure scambiandoci dichiarazioni non vincolanti relative alle nostre intenzioni, allora la regolarità risultante sarebbe stata subito una convenzione perfetta.

Per renderci conto del modo in cui una convenzione potrebbe prendere origine da un accordo o da qualcos'altro, prendiamo come esempio una con-

venzione che non esiste ancora: una convenzione tra i logici che stabilisca una notazione standard. Attualmente non esiste una convenzione del genere. Siccome vengono usate diverse notazioni, ciascuno si sente libero di indulgere alle proprie preferenze. Ma una tale convenzione potrebbe esistere. Se una delle notazioni attuali fosse usata altrettanto uniformemente che la notazione aritmetica standard, tutti se ne servirebbero. Nessuno continuerebbe a far uso della propria notazione favorita se un'altra fosse quella standard. (Una notazione inusitata sarebbe più fastidiosa per lettori abituati a una notazione standard che per lettori abituati al nostro caos attuale.)

Immaginiamo che la diversità di notazioni diventi più fastidiosa di quanto è oggi. Ogni autore inventa la propria notazione; alcuni la adottano, altri la imparano ma scelgono di non farne uso, e altri non riescono neppure a leggerla. L'*Association for Symbolic Logic* potrebbe allora tenere una seduta in cui venga discusso il problema; ciascuno dei presenti esprime la speranza che venga stabilita una notazione standard; e la notazione dei *Principia* viene scelta con una votazione. In questo caso, i partecipanti alla seduta considerano se stessi, e si considerano reciprocamente, come coloro che hanno espresso l'intenzione di usare d'ora in poi la notazione dei *Principia* se lo farà anche la maggior parte degli altri logici. Il loro scambio di dichiarazioni circa le proprie intenzioni, il quale avviene in una seduta a partecipazione diretta, costituisce un accordo sufficiente a dare origine a una convenzione.

Ma la stessa convenzione potrebbe invece aver origine in altri modi. Immaginiamo qualche possibilità.

Alcuni eminenti logici - per esempio, i redattori del *Journal of Symbolic Logic* - potrebbero pubblicare sul JSL una dichiarazione congiunta secondo la quale va posto un rimedio alla fastidiosa diversità di notazioni, per cui, a loro avviso, d'ora in poi andrebbe usata esclusivamente la notazione dei *Principia*. Siccome tra i logici è conoscenza comune il fatto che essi leggono tutti il JSL, diventerebbe conoscenza comune anche il fatto che la maggior parte di loro ha letto la dichiarazione. E, se era conoscenza comune tra i logici il fatto che la maggior parte di loro sarebbe stata propensa a seguire tale indicazione se l'avessero fatto gli altri, allora sarebbe conoscenza comune il fatto che la maggior parte dei logici ha una propensione molto maggiore a far uso della notazione dei *Principia*. Ammesso un interesse sufficiente a conformarsi a un qualunque standard, ne potrebbe derivare la convenzione di far uso della notazione dei *Principia*. In questo caso soltanto alcune persone svolgono un ruolo attivo nel dare origine alla nuova convenzione; gli altri sono un uditorio sensibile.

Oppure potrebbe non esserci alcun tentativo di creare una nuova convenzione. Parecchi logici, disturbati dalla proliferazione delle notazioni e

preoccupati dell'intelligibilità dei loro scritti, potrebbero autonomamente decidere di passare dalle loro notazioni preferite a quella che sembra la più nota, per esempio quella dei *Principia*. Il maggior uso della notazione dei *Principia* che ne deriverebbe potrebbe apparire come l'inizio di una tendenza, per cui gli altri sarebbero inclini a cambiare notazione e si aspetterebbero reciprocamente di cambiare notazione. Il risultato sarebbe di nuovo una convenzione.

Oppure potrebbe accadere che, per nessuna ragione particolare, ci fosse una grande diffusione di opere scritte nella notazione dei *Principia*. Per pura coincidenza, potrebbe succedere che molti dei logici che preferiscono la notazione dei *Principia* si mettessero a pubblicare contemporaneamente. Anche in questo caso si manifesterebbe una sensibile tendenza, che gli altri potrebbero seguire.

In breve, certe condizioni (la conoscenza comune di una preferenza generale a far uso di una notazione sufficientemente nota, più la conoscenza comune del fatto che i logici sanno dire quando vengono usate le varie notazioni) tendono in qualche modo a ampliare qualunque fluttuazione delle aspettative e delle propensioni dei logici relative alla loro scelta di notazione. La convenzione si produce quando una fluttuazione sufficientemente grande incontra forze amplificatrici sufficientemente intense. Ammesso che abbia una certa consistenza, l'origine della fluttuazione non è importante. Non importa se era stata prodotta con l'intenzione di dare origine a una convenzione o se si era manifestata in tutta la popolazione o in una parte di essa.

Ammesso che l'accordo esplicito è solo una delle numerose origini possibili della convenzione, possiamo ancora chiederci se esso gode di prerogative particolari. È forse vero che tutte le convenzioni *potrebbero* nascere per accordo? Ecco tre controargomentazioni.

Innanzitutto, ricordiamo l'esempio dei rematori di Hume, una regolarità convenzionale che non possiamo descrivere. Non possiamo descriverla in pratica; se fosse possibile in linea di principio, occorrerebbero più tempo e più strumenti di misurazione di quanti i rematori possano avere a disposizione. Ma questa argomentazione arriva soltanto a dimostrare che la loro convenzione non potrebbe trarre origine da un accordo che sia un semplice scambio verbale. Essi possono benissimo accordarsi di remare *così*, specificando il ritmo di vogata per mezzo dell'esibizione di esso.

Inoltre, supponiamo che una convenzione produca coordinazione per la realizzazione di un fine che sia in qualche modo compromesso dall'atto stesso dell'accordarsi. Quella convenzione non potrebbe quindi nascere per accordo, in quanto un accordo distruggerebbe l'aspetto essenziale del conformarsi

alla convenzione. Prendiamo due persone che trovano conveniente mantenere un'apparenza di ostilità (per esempio, i leader di due partiti politici rivali). Al fine di non trovarsi pubblicamente d'accordo - con imbarazzo di entrambi essi potrebbero servirsi di una convenzione che specifichi che tipo di opinioni debbano professare su qualsiasi argomento. Ma non possono stabilire una convenzione accordandosi. Ciò infatti distruggerebbe la loro apparenza di ostilità rivelando il loro comune interesse a conservarla, e li priverebbe di ogni possibilità di accordo. (È vero che potrebbero accordarsi segretamente. Ma la segretezza non li aiuterebbe se - modificando l'esempio - supponiamo che quella facciata ostile essi la presentino innanzitutto *a se stessi*. Se è così, però, il controesempio risulta sospetto, perché gli agenti ingannano se stessi. Essi credono nella propria apparente ostilità e nello stesso tempo prendono delle precauzioni per non incontrare la prova che distruggerebbe la loro illusione. Perciò non possono essere senz'altro trattati alla stregua di agenti razionali con credenze coerenti.)

Infine, consideriamo questa argomentazione, proposta da Quine e da altri.<sup>2</sup> La convenzione di linguaggio che viene stabilita per prima non potrebbe nascere da un accordo stipulato in una lingua regolata da convenzioni. Così, anche se *qualunque* convenzione di linguaggio potesse derivare da un accordo di questo tipo, non lo potrebbero tutte. (Questa argomentazione è quindi diversa dalle precedenti, che erano intese a mostrare come convenzioni particolari non potrebbero derivare da un accordo.) Propongo questa risposta: un accordo bastate a creare una convenzione non deve necessariamente essere una transazione implicante il linguaggio o una qualunque altra attività convenzionale. Tutto quello che occorre è uno scambio di manifestazioni di una propensione a conformarsi a una regolarità. Queste manifestazioni potrebbero essere semplici esibizioni di un'azione conforme in diverse situazioni appropriate, compiute durante un convegno a partecipazione diretta per dare origine a una convenzione. Tale scambio di esibizioni si potrebbe chiamare "accordo" senza forzare troppo il termine.

Ritengo che nessuna delle tre argomentazioni sia decisiva. Assumendo la parola "accordo" in un'accezione ampia, tutte le convenzioni potrebbero forse, in linea di principio, trarre origine da un accordo. Quello che è chiaro è che non è necessariamente così. E spesso non è così: il capitolo

---

<sup>2</sup>Da Quine in "Truth by Convention"; da Bertrand Russell in *The Analysis of Mind*, Allen and Unwin, London, 1921, p. 190; da William Alston in *Philosophy of Language*, Prentice-Hall, Englewood Cliffs, 1964, p. 57 [trad. it. di G. Sandri e R. Simili, *Filosofia del linguaggio*, Il Mulino, Bologna, 1971, p. 96 (N.d.T.)]; tutti e tre ritengono che si tratti di un'argomentazione contro l'esistenza di convenzioni di linguaggio, in quanto pensano che le convenzioni che meritano questo nome debbano essere stabilite per accordo.

1.5 dovrebbe suggerire esempi familiari di convenzioni che hanno un'origine diversa. Le convenzioni sono come il fuoco: in condizioni favorevoli, una concentrazione sufficiente di calore si propaga e si perpetua. La natura del fuoco non dipende dalla fonte di calore iniziale. Può darsi che i fiammiferi siano gli strumenti migliori per produrre fuoco, ma non c'è alcun motivo di ritenere che il fuoco prodotto diversamente abbia titoli minori per essere considerato un vero e proprio fuoco.

### 3.2 Contratti sociali

Sembra (nonostante alcune deboli obiezioni) che una convenzione sia una regolarità di comportamento che vale *come se* fosse conseguenza dell'accordo di comportarsi così, in virtù di una preferenza generale per la conformità generale a quella regolarità. Ora, questo è esattamente il modo in cui si potrebbe descrivere un contratto sociale, una volta che si sia accettata l'artificiosità del trattare la stipulazione originaria del contratto come la drammatizzazione fittizia dei nostri attuali motivi di conformarci. Il mio concetto di convenzione non si riduce forse al familiare concetto di contratto sociale, quale l'abbiamo ereditato da Hobbes, Locke e Rousseau, demitologizzato e applicato a questioni diverse dall'obbedienza politica e dalla solidarietà sociale?

No. Il concetto di contratto sociale, come lo concepisco io, è diverso in linea di principio da quello di convenzione (per quanto se ne diano descrizioni, come la precedente, abbastanza sommarie da oscurare ogni differenza). Né le estensioni di questi concetti coincidono, sebbene siano in gran parte sovrapposte.

Propongo di definire grosso modo il *contratto sociale* come una regolarità R del comportamento dei membri di una popolazione P, quando essi sono agenti in una situazione S, tale che è vero, ed è conoscenza comune, che:

Qualunque membro di P implicato in S si comporta in conformità a R.

Ciascun membro di P preferisce lo stato di conformità generale a R (da parte di membri di P in S) a un certo stato contestualmente definito di non-conformità generale a R, chiamato *stato di natura* relativo al contratto sociale R.

Lo stato di natura non è solo uno stato di non-conformità generale a R (da parte di membri di P in S); esso viene in qualche modo specificato.<sup>3</sup> Lo

---

<sup>3</sup>Cfr. R.P. Wolff, "A Refutation of Rawls' Theorem on Justice", *Journal of Philosophy*,

stato di natura relativo al contratto sociale di Hobbes (a causa del quale costituiamo un leviatano per mezzo dell'obbedienza regolare a un sovrano) è concepito come anarchia e guerra di tutti contro tutti. Non è l'anarchia pacifica, né l'esistenza di un leviatano sotto un altro sovrano, sebbene anche questi siano stati di non conformità generale al nostro attuale contratto sociale. L'anarchia pacifica non va bene perché Hobbes la ritiene instabile; l'esistenza di un leviatano sotto qualche altro sovrano non va bene perché è di una specie troppo simile al contratto sociale in questione.

Lo stato di natura relativo a R è uno stato di non conformità generale a R; uno stato di non conformità generale a altre regolarità di un tipo simile a R; e uno stato in cui nessuno fa molto assegnamento su previste regolarità del comportamento degli altri. Nessuno ci perde troppo se le sue aspettative circa i propri vicini si dimostrano sbagliate. Si tratta di uno stato particolarmente stabile. Ed è uno stato in cui potremmo venire a trovarci se dovessimo partire in qualche modo da zero senza forti aspettative reciproche (di qui la plausibilità *a priori* del mito secondo il quale i popoli primitivi vivono in uno stato di natura).

Si osservi che, siccome lo stato di natura relativo a R è possibile, ne segue che i membri di P non preferiscono incondizionatamente di conformarsi a R - il che, se non ne fosse stato implicato, avrebbe dovuto essere inserito nella definizione di contratto sociale.

La mia definizione di contratto sociale si manteneva il più possibile parallela a quella di convenzione, in modo da far vedere dove sta la differenza: nella natura della preferenza generale per la conformità generale. Preferire qualcosa significa preferirlo *a* qualcos'altro, e il secondo termine della preferenza non è lo stesso. Per la convenzione noi richiediamo che ciascun agente preferisca la conformità generale alla conformità da parte di tutti tranne lui, ignorando le sue preferenze relative agli stati di non-conformità generale. Per il contratto sociale richiediamo che ciascun agente preferisca la conformità generale a un certo stato di non-conformità generale, ignorando le sue preferenze relative alla conformità da parte di tutti tranne lui.

Consideriamo la regolarità R dell'obbedire a un sovrano. Vediamo come R potrebbe essere una convenzione, un contratto sociale, entrambi, o nessuno dei due. Supponiamo che lo *status quo* sia il seguente: noi obbediamo quasi sempre agli ordini del sovrano:

---

63 (1966), pp. 179-190. Wolff, replicando al contrattualismo ideale proposto da John Rawls in "Justice as Fairness", *Philosophical Review*, 67 (1958), pp. 164-194, obietta che Rawls non ha indicato alcun modo di identificare la "linea di base" - cioè lo stato di natura - con cui confrontare un contratto sociale.

- (1) di evitare di rubarci reciprocamente i nostri averi;
- (2) di consegnare tutto quello che risparmiamo per permettere al sovrano una vita lussuosa;
- (3) di collaborare a catturare e punire chiunque trasgredisca a uno di questi tre ordini.

(Se qualcuno è urtato da questa caricatura della società politica, può considerarla relativa alla malavita, ai gangster, o alle nazioni.) Ciascuno di noi ha una qualche gerarchia preferenziale nei confronti dei seguenti tre stati, ciascuno con i suoi vantaggi e i suoi svantaggi:

- (SQ)Lo status quo. Sono protetto dai miei vicini grazie al potere che il sovrano ha di far rispettare il suo primo comando. Sono al sicuro dal potere poliziesco del sovrano perché non cerco di disobbedire. Ma sono povero perché contribuisco a mantenere il sovrano e non mi impadronisco degli averi degli altri.
- (SN)Lo stato di natura. Nessuno impone l'obbedienza generale, per cui non devo temere alcun potere poliziesco organizzato. Non contribuisco a mantenere un sovrano, perché non ce ne sono. Ma devo impegnarmi a proteggere i miei averi dai vicini, e vivo nel timore di non riuscirci.
- (DI)La disobbedienza individuale. I miei beni sono protetti dai vicini. Posso diventare ricco perché qualche volta mi impadronisco degli averi dei miei vicini e non verso tutto quello che devo per mantenere il sovrano. Ma vivo nel timore che il sovrano, con l'aiuto dei miei vicini, mi catturi e mi punisca.

Lasciando da parte l'indifferenza, ciascuno di noi deve avere una di queste sei gerarchie preferenziali, non necessariamente la stessa per ciascuno:

<i>SN</i>	<i>SQ</i>	<i>SQ</i>	<i>DI</i>	<i>SN</i>	<i>DI</i>
<i>SQ</i>	<i>DI</i>	<i>SN</i>	<i>SQ</i>	<i>DI</i>	<i>SN</i>
<i>DI</i>	<i>SN</i>	<i>DI</i>	<i>SN</i>	<i>SQ</i>	<i>SQ</i>

(dove lo stato maggiormente preferito è in alto, quello meno preferito in basso). Le preferenze di ognuno dipenderanno dal suo temperamento; da quello che pensa del carattere dei suoi vicini e del sovrano; dalla sua valutazione della probabilità che gli altri imitino la sua disobbedienza; dalla sua valutazione della propria capacità di guadagnarsi da vivere, di difendersi nello stato di natura, o di sfuggire alla punizione della disobbedienza; e dal

valore che assegna alla sicurezza, alla ricchezza, alla liberazione dal dovere dell'autodifesa, al benessere dei suoi vicini, al benessere del sovrano, alla giustizia, alla pace e alla fiducia.

Se R è una convenzione, le uniche gerarchie preferenziali che noi abbiamo (tranne una minoranza trascurabile) sono le tre in cui SQ è posto sopra DI, in modo che ciascuno preferisce conformarsi a R subordinatamente alla conformità da parte degli altri:

$$\begin{array}{ccc} SN & SQ & SQ \\ SQ & DI & SN \\ DI & SN & DI \end{array}$$

Non importa dove figura SN nella gerarchia di qualcuno.

Se R è un contratto sociale, le uniche gerarchie preferenziali che noi (tranne una minoranza trascurabile) abbiamo sono le tre in cui SQ è posto sopra SN, in modo che ciascuno beneficia della conformità generale a R:

$$\begin{array}{ccc} SQ & SQ & DI \\ DI & SN & SQ \\ SN & DI & SN \end{array}$$

Non importa dove figura DI nella gerarchia di qualcuno.

Se R è tanto un contratto sociale che una convenzione, le uniche gerarchie che noi (tranne una minoranza trascurabile) abbiamo sono le due in cui SQ è preferito a entrambe le alternative:

$$\begin{array}{cc} SQ & SQ \\ DI & SN \\ SN & DI \end{array}$$

Se R è una convenzione ma non un contratto sociale, noi tutti o una parte di noi ha la gerarchia preferenziale:

$$\begin{array}{c} SN \\ SQ \\ LD \end{array}$$

Costoro si trovano in un vicolo cieco. Vogliono che la convenzione sia abbandonata. Ma nessuno osa abbandonarla se non può essere sicuro che molti altri l'abbandonino con lui. Per prendere il caso estremo: se *tutti* noi preferiamo lo stato di natura allo status quo, la convenzione deve interamente la propria sopravvivenza alla difficoltà di organizzare una disobbedienza *concertata* ai comandi del sovrano. (Se al sovrano sta a cuore la propria posizione, farà bene a render noto un quarto ordine che proibisce gli sforzi di organizzazione in questo senso.) Tutte le convenzioni sono metastabili, ma questa è destinata a essere meno stabile di quasi tutte le altre, in quanto abbiamo l'incentivo a metterci insieme per modificarla, se ne siamo capaci. La convenzione di non ammettere che l'imperatore non avesse vestiti non era un contratto sociale. Ciascuno desiderava infrangerla, ma soltanto il fanciullino osò infrangerla *da solo*. Quando lo fece, la convenzione crollò e fu restaurato lo stato di natura.

Se R è un contratto sociale ma non una convenzione, noi tutti o una parte di noi ha la gerarchia preferenziale:

*DI*  
*SQ*  
*SN*

Ci si può chiedere come mai costoro continuino a conformarsi. Infatti, lasciandosi sfuggire le occasioni di trarre profitto dalla disobbedienza agli ordini del sovrano, essi si comportano inevitabilmente in contrasto con le proprie preferenze.

Se concepiamo le preferenze di una persona come il risultato di tutti gli impulsi più o meno durevoli che entrano nella determinazione delle sue scelte, allora l'azione che contrasta regolarmente con la preferenza è a malapena possibile. Dovrebbe essere dovuta a impulsi transitori, diversi in occasioni diverse. Ed è difficile immaginare come possa diventare conoscenza comune che delle persone si comporterebbero regolarmente contro le proprie preferenze, in quanto l'azione in contrasto con la preferenza è intrinsecamente eccezionale. Così, in base a questo concetto di preferenza, che è quello che ci è più familiare, è quasi impossibile che un contratto sociale non sia una convenzione per il fatto che alcuni preferiscono non conformarsi sebbene gli altri si conformino.

A volte, tuttavia, interpretiamo la preferenza in una accezione più limitata, ossia come il risultato di quegli impulsi determinanti la scelta che sono *diversi dal* senso del dovere. Gli obblighi morali accettati da noi possono

prevalere, e invero lo fanno regolarmente, sulle nostre preferenze in questo senso stretto. Dunque potremmo dire di una persona che le sue preferenze sono

*DI*  
*SQ*  
*SN*

ma che egli obbedisce ai comandi del sovrano in contrasto con le proprie preferenze perché si ritiene moralmente obbligata a comportarsi così. Può darsi che quella persona pensi, con Locke, che, quando per la prima volta ha trascurato l'occasione di andarsene dal paese, ha tacitamente promesso ai suoi compagni di obbedire se essi obbediscono; oppure può darsi che creda, con John Rawls, di essere obbligato a comportarsi correttamente per ricambiare i benefici che gli sono derivati, e che ha prontamente accettato, dall'obbedienza altrui.<sup>4</sup> In ogni caso, il suo obbligo deriva dal fatto che preferisce lo *status quo* allo stato di natura. Quindi, per qualunque contratto sociale, le condizioni di tali obblighi sono presenti per ciascuno. Se ciascuno riconoscerà tali obblighi, ciascuno rispetterà il contratto sociale, sia questa la sua preferenza o meno. Il contratto sociale persisterà, e può darsi che ciò sia conoscenza comune. Ma, se le preferenze menzionate nella definizione di convenzione vengono intese nell'accezione ristretta, non è una convenzione.

Se ora torniamo alla nostra nozione comune, e più ampia, di preferenza, rimane vero che molti contratti sociali saranno tenuti in vita dagli obblighi morali dell'accordo tacito o del comportamento corretto, quali vengono riconosciuti dagli agenti implicati. Ma questi obblighi accettati saranno considerati come componenti delle preferenze, non come impulsi autonomi determinanti le nostre scelte. Siccome accettiamo questi obblighi, le nostre gerarchie preferenziali saranno:

*SQ*     *SQ*  
*DI*   *o*   *SN*  
*SN*     *DI*

anziché

---

<sup>4</sup>Ma il contenuto del nostro contratto sociale presentato a scopo illustrativo non è il contenuto che avevano in mente Locke o Rawls; è piuttosto il contratto semplice e disperato del *Leviathan* di Hobbes. Non vedo alcuna ragione per cui gli uomini non dovrebbero aderire al contratto di Hobbes per i motivi di Locke o Rawls.

*DI*  
*SQ*  
*SN*

quali avrebbero potuto essere altrimenti (per alcuni di noi). Così il nostro contratto sociale è, in ultima analisi, una convenzione. Ma è una convenzione a causa della modificazione delle nostre preferenze operata dagli obblighi, e questi obblighi esistono perché vi è un contratto sociale.

Così, la possibilità di un contratto sociale che non sia una convenzione (nel modo fin qui considerato) è problematica; dipende dall'adozione del meno comune dei nostri due concetti di preferenza, quello per cui la preferenza è opposta all'obbligo.

C'è un modo del tutto diverso, meno problematico, in cui un contratto sociale può non essere una convenzione. Teniamo presente che avevamo richiesto che una convenzione fosse alternativa a diverse altre, mentre il contratto sociale non deve avere altra alternativa che lo stato di natura. Lo stato di natura non deve necessariamente essere uno stato in cui raggiungiamo degli equilibri di coordinazione conformandoci a una regolarità. Lo stato di natura di Hobbes non è certamente tale: nella guerra di tutti contro tutti, le battaglie non solo non mettono capo a equilibri di coordinazione,<sup>5</sup> ma neppure a equilibri (in quanto il perdente desidererà aver adottato una strategia diversa). Un contratto sociale può quindi non essere una convenzione per mancanza di alternative, anche se è una regolarità per mezzo della quale raggiungiamo degli equilibri di coordinazione e in cui, perciò, ciascuno preferisce conformarsi se lo fanno gli altri.

Se ci troviamo ripetutamente implicati in una certa situazione (che non sia un problema di coordinazione) illustrata dalla matrice delle vincite della Figura 30, e ci comportiamo regolarmente tutti in modo da raggiungere l'equilibrio di coordinazione  $\langle R1, C1 \rangle$ , la nostra regolarità è un contratto sociale, ed è una regolarità per mezzo della quale raggiungiamo degli equilibri di coordinazione, ma non è una convenzione. Infatti c'è soltanto un altro stato stabile, lo stato di natura: quello in cui scegliamo ogni volta a caso tra  $R2$  e  $R3$  (o tra  $C2$  e  $C3$ ). E i risultati che otteniamo nello stato di natura -  $\langle R2, C2 \rangle$ ,  $\langle R2, C3 \rangle$ ,  $\langle R3, C2 \rangle$  e  $\langle R3, C3 \rangle$  - non sono equilibri di

---

<sup>5</sup>Nell'esempio di Rousseau della caccia al cervo, d'altra parte, lo stato di natura relativo al contratto sociale consistente nell'aiutarsi tutti a cacciare insieme il cervo è lo stato in cui catturiamo tutti dei conigli per conto nostro; e questo stato di natura è uno stato in cui raggiungiamo degli equilibri di coordinazione.

	C1	C2	C3
R1	1	0,5	0,5
R2	-0,5	1	-1
R3	-0,5	-1	1

Figura 30

coordinazione, anzi, neppure equilibri. (Raggiungiamo, in effetti, una combinazione equilibrata di *strategie miste*, ma si tratta di un equilibrio solo in un'accezione ampia del termine, non ancora di un equilibrio di coordinazione.) Così lo stato di natura non è una convenzione, e neppure lo è la nostra regolarità consistente nello scegliere R1 e C1.

Per fare un esempio, supponiamo di essere sempre le stesse dieci persone a recarci ogni venerdì in un ristorante cinese dove ci viene servito, tra l'altro, un piatto di venti gamberetti fritti. A noi piacerebbe mangiarne tre o quattro a testa, ma per non spiacciare agli altri ciascuno è disposto a limitarsi a due se e solo se anche ciascun altro si limita. Nessuno si cura di trattenersi se non quando è sicuro che in tal modo permette a ciascuno di avere la propria parte. Ci sono quindi due stati stabili: un contratto sociale, in virtù del quale ciascuno ne prende due, o lo stato di natura, una lotta in cui i primi arrivati prendono tutti i gamberetti che vogliono finché non ne rimangono più. Lo stato di natura non è uno stato in cui raggiungiamo degli equilibri di coordinazione, e quindi il contratto sociale non è una convenzione, sebbene ciascuno preferisca conformarsi se lo fanno gli altri. Ed è giusto che sia così: questo contratto sociale sembra mancare dell'arbitrarietà caratteristica della convenzione.

Infine, un contratto sociale potrebbe non essere una convenzione, o vi-

ceversa, in un altro modo ancora: le conoscenze comuni richieste nella definizione di convenzione non sono le stesse di quelle richieste nella definizione di contratto sociale. Quindi, un'opportuna mancanza di conoscenza comune potrebbe impedire a una regolarità di essere una convenzione ma non di essere un contratto sociale, o viceversa. È difficile concepire come potrebbe verificarsi tale mancanza di conoscenza comune discriminante; in generale, le condizioni che determinano la conoscenza comune di uno qualunque dei fatti di un certo esempio determinano la conoscenza comune di tutti i fatti di quell'esempio.

### 3.3 Norme

La definizione di convenzione che ho dato non conteneva termini normativi: “dovere”, “bene” e simili. Né abbiamo ragione di attenderci che in qualunque definizione equivalente compaiano necessariamente dei termini normativi. Così, lo stesso termine “convenzione” non è, in base alla mia analisi, un termine normativo.

Ciononostante, le convenzioni possono essere una specie di norme: regolarità alle quali crediamo che ci si debba conformare. Ciò è quanto sosterrò. Ci sono certe conseguenze probabili implicate dal fatto che un'azione si conformi a una convenzione (qualunque sia l'azione e qualunque sia la convenzione) e essi sono presunti motivi per cui, in base alle nostre opinioni correnti, un'azione dovrebbe essere compiuta.

Supponiamo che R sia una convenzione, in una popolazione P, relativa al comportamento in una situazione S. Allora questa supposizione, in base alla definizione di convenzione, e indipendentemente da quello che possono essere R, P e S, fa sì che sia probabile quanto segue:

- (1) La maggior parte dei membri di P implicati con me in una situazione S si conformerà a R.
- (2) Io preferisco conformarmi, se la maggior parte dei membri di P implicati con me in S si conforma.
- (3) La maggior parte dei membri di P implicati con me in S si aspetta, a ragione, che io mi conformi a R.
- (4) La maggior parte dei membri di P implicati con me in S preferisce che, se la maggior parte di essi si conforma, anch'io mi conformi.
- (5) Io ho motivo di credere che (1)-(4) valgano.

Se avessimo supposto che R è una convenzione di gradi 1, 1, 1, 1, 1, 1, allora questa supposizione, insieme alla supposizione che io fossi un mem-

bro di P in S, avrebbe implicato (1)-(5), addirittura rafforzati sostituendo dappertutto “la maggior parte” con “tutti”. Ma, siccome abbiamo scelto di permettere che ci siano convenzioni non perfettamente convenzionali, dobbiamo badare sia a formulare (1)-(5) in modo da permettere delle eccezioni, sia a permettere delle eccezioni anche a (1)-(5). Ma (1)-(5) devono valere nella *maggior parte* dei casi in cui si decide se conformarsi a una convenzione. Così in qualunque caso dato - tranne quando è provato che si tratta di un caso eccezionale - ciascuno dei fatti (1)-(5) probabilmente vale; per cui, probabilmente (anche se la probabilità è minore) essi valgono tutti.

E se essi valgono tutti in un qualunque caso, allora valgono anche questi:

- (6) Io ho motivo di credere che il mio conformarmi sarebbe rispondente alle mie preferenze.
- (7) Io ho motivo di credere che il mio conformarmi sarebbe rispondente alle preferenze della maggior parte degli altri membri di P implicati con me in S; e che essi hanno motivo di aspettarsi che io mi conformi.

E (6) e (7), se veri, sono presunti motivi per cui dovrei conformarmi. Infatti noi presumiamo, a parità di condizioni, che si debba fare ciò che risponde alle proprie preferenze. E presumiamo, a parità di condizioni, che si debba fare ciò che risponde alle preferenze degli altri, specialmente quando questi possono ragionevolmente aspettarsi che ci si comporti così. Per una qualunque azione conforme a una qualunque convenzione, dunque, riconosceremmo questi due motivi (probabili e presunti per cui dovrebbe essere compiuta. Non riconosceremmo, a quanto ne so, motivi simili per cui non dovrebbe essere compiuta. È questo che voglio dire quando dico che le convenzioni sono una specie di norme.

Naturalmente, per qualunque azione particolare conforme a una convenzione, possono esserci moltissime ragioni per cui dovrebbe o non dovrebbe essere compiuta. (Potrebbero anche esserci motivi per credere che quello in questione è un caso eccezionale in cui o (6) o (7) non vale.) La convenzione potrebbe essere stata adottata per mezzo di uno scambio di promesse, per cui l'azione conforme dovrebbe essere compiuta per mantenere la propria promessa. Oppure potrebbe essere una convenzione che è anche un contratto sociale, per cui l'azione conforme dovrebbe essere compiuta per ricambiare il beneficio che si trae dalla conformità altrui. D'altra parte, potrebbe essere un patto tra oligopolisti per fissare i prezzi, o tra ladri per lavorare insieme, per cui l'azione conforme non dovrebbe essere compiuta perché non di interesse pubblico. Come sempre, i vari motivi presunti per cui un'azione

dovrebbe o non dovrebbe essere compiuta vengono soppesati; può benissimo darsi che (6) e (7) vengano vanificati. La loro importanza non sta nel fatto che sono considerazioni particolarmente importanti. Caso mai, sta nel fatto che sono particolarmente generali: per qualsivoglia convenzione, esse devono influire sulla decisione se conformarsi ad essa o meno.

Una convenzione è, per definizione, una norma alla quale si suppone che ci si debba conformare. Adesso cercherò di dimostrare che, per definizione, è anche una norma socialmente imposta: ci si aspetta che una persona si conformi, e la mancata conformazione tende a suscitare risposte sfavorevoli da parte degli altri. Infatti, se R è una convenzione di una popolazione P relativa al comportamento in una situazione S, e io sono un membro di P in S, allora probabilmente (per la definizione di convenzione, e indipendentemente da quello che possono essere R, P e S):

- (8) La maggior parte degli altri membri di P implicati con me in S si aspetta che io mi conformi.
- (9) La maggior parte degli altri membri di P implicati con me in S ha motivo di credere che le condizioni (1)-(5) valgano.

E, ogniqualvolta vale (9), valgono anche:

- (10) La maggior parte dei membri di P implicati con me in S ha motivo di credere che il mio conformarmi risponderebbe alle mie preferenze.
- (11) La maggior parte dei membri di P implicati con me in S ha motivo di credere che io abbia motivo di credere sia che il mio conformarmi risponderebbe alle loro preferenze sia che essi abbiano motivo di aspettarsi che io mi conformi.

Così, se essi si accorgono che io non mi conformo, non ho soltanto deluso le loro aspettative; essi saranno probabilmente in condizione di dedurre che mi sono coscientemente comportato in contrasto con le mie preferenze, e in contrasto con le loro preferenze e con le loro ragionevoli aspettative. Saranno sorpresi, e tenderanno a spiegare la mia condotta in modo disonorevole per me. La cattiva opinione che essi si fanno di me, i loro rimproveri, le loro punizioni e la loro mancanza di fiducia sono le risposte sfavorevoli che ho suscitato con il non essermi conformato alla convenzione.

Consideriamo il nostro esempio di convenzione secondo cui; quando una telefonata è interrotta colui che ha chiamato ri chiama, mentre chi è stato chiamato aspetta. Supponiamo che io non mi conformi, e cioè che aspetti quando ho chiamato oppure richiami quando sono stato chiamato. Il mio

*partner* sa quello che ho fatto. Sa che avrei dovuto sapere che comportandomi in quel modo non avrei ristabilito la comunicazione. Può supporre che io l'abbia fatto senza pensarci; o che sono troppo ottuso per imparare la convenzione; o che ero comunque stanco di parlare con lui e non volevo ristabilire la comunicazione: oppure che mi aspettavo che *lui* fosse colpevole di una di queste manchevolezze e che ho violato la convenzione per neutralizzare la violazione che mi aspettavo da lui. Qualunque cosa pensi, va a scapito dell'opinione che ha di me. Lo stesso vale per il modo in cui probabilmente mi tratterà in futuro. Sono conseguenze negative, e il mio interesse a evitarle rafforza la mia preferenza condizionale per il rispetto della convenzione.

### 3.4 Regole

Certamente chiameremmo regole molte convenzioni. Per esempio, quelle presentate nel capitolo I.5 quali soluzioni degli undici problemi di coordinazione campione avrebbero potuto essere chiamate in modo naturale regole, anche se, probabilmente, con una restrizione: regole “tacite”, regole “informali”, regole “non scritte”, o simili. Ma certamente non tutte le cosiddette regole sono convenzioni. Consideriamo diversi tipi di controesempi.

A volte vengono chiamate regole delle semplici generalizzazioni, delle leggi di natura o anche delle verità matematiche. Queste regole possono non aver nulla a che fare con la condotta degli agenti umani, tranne per il fatto che gli agenti umani possono trarre beneficio dal prenderle in considerazione. Per esempio, in un libro di cucina troviamo questo passo:

Ecco una regola fondamentale che ammette pochissime eccezioni: tutte le carni sono più tenere e succose se cotte a bassa anziché ad alta temperatura.<sup>6</sup>

E questo è un teorema che un testo di algebra chiama “Regola di Cartesio dei segni”:

Un'equazione  $f(x) = 0$  non può avere più radici positive di quanti siano i cambiamenti di segno in  $f(x)$ , e non può avere più radici negative di quanti siano i cambiamenti di segno in  $f(-x)$ .<sup>7</sup>

---

<sup>6</sup>B.B. McLean e T.W. Campbell, *The Meat and Poultry Cookbook*, Pocket Books, New York, 1960, p. 19.

<sup>7</sup>H.S. Hall e S.R. Knight, *Higher Algebra*, Macmillan, London, 1960, p. 459.

Altre cosiddette regole sono massime strategiche, imperativi ipotetici che dicono quello che un agente umano dovrebbe fare per conseguire un qualche scopo. Queste regole esprimono generalizzazioni relative alla tendenza di certe azioni a realizzare certi scopi. Ci vengono per esempio date queste “regole generali per l’uso di un insetticida”:

Trattate qualsiasi insetticida domestico, quale che sia la sua etichetta, come un veleno. Non spruzzate mai gli insetticidi sulle pareti della nursery, sul recinto e sul letto dei bambini, sui giocattoli o nei luoghi in cui si trovano dei neonati. Comprate insetticidi solo quando ne avete bisogno. Metteteli in un armadietto chiuso a chiave. Dopo che gli insetti sono stati debellati, gettate tutto l’insetticida rimasto in fondo alla pattumiera.<sup>8</sup>

E questa è la “regola per ridurre un numero decimale periodico a una frazione ordinaria” (un imperativo ipotetico basato su un teorema):

Al numeratore, sottrarre il numero intero consistente delle cifre non periodiche dal numero intero consistente delle cifre non periodiche e periodiche; al denominatore prendere un numero consistente di tanti nove quante sono le cifre periodiche seguiti da tanti zeri quante sono le cifre non periodiche.<sup>9</sup>

Ecco delle regole per educare il gusto:

La prima regola per l’intenditore di vini è fidarsi del proprio palato, ossia credere nella sua capacità fisica registrare tante sensazioni quante ne registra il palato di chiunque altro... La seconda regola importante per l’intenditore è avere il coraggio di cambiare opinione. Via via che l’esperienza aumenta e la percezione si fa più acuta, il suo gusto cambierà certamente e i vini che in un primo tempo gli piacevano possono adesso sembrargli banali o non piacergli decisamente.<sup>10</sup>

Altre regole sono imperativi ipotetici rafforzati da una codificazione autorevole e fatti rispettare per mezzo di sanzioni. Questa regola potrebbe comparire su un cartello in uno stabilimento chimico:

Ai dipendenti è vietato fumare nel raggio di 100 metri dai contenitori di acetone; ogni infrazione sarà considerata motivo di licenziamento immediato.

---

<sup>8</sup> *Consumer Reports*, 30 (1965), no. 12, p. 156.

<sup>9</sup> Hall e Knight, *Higher Algebra*, p. 43.

<sup>10</sup> Allan Sichel, *The Penguin Book of Wines* (Penguin Books, Baltimore, 1965), p. 22.

Non fumare nei pressi di un contenitore di acetone concorda sempre con l'interesse che ogni lavoratore ha a evitare incendi, indipendentemente dal fatto che i suoi compagni fumino o meno vicino ai contenitori di acetone e dal fatto che la direzione abbia emanato una regola per vietarlo. La regola ufficiale ricorda ai dipendenti questo imperativo ipotetico. Essa li informa anche del fatto che la prassi seguita dalla direzione è di silurare chiunque sia sorpreso a fumare vicino ai contenitori di acetone; è una minaccia o un preavviso<sup>11</sup> inteso a scoraggiarli dal farlo un'altra volta.

Altre regole sono minacce o preavvisi sanciti da qualche autorità o governo per controllare il comportamento di una classe di persone, contro le loro stesse preferenze. In un campo di concentramento potrebbe esserci la regola che i prigionieri non devono riunirsi in gruppi di più di sei persone, a rischio di essere puniti con dieci giorni a pane e acqua. Una *gang* locale specializzata in estorsioni potrebbe imporre la regola secondo cui le tavole calde devono affittare un *flipper* per ogni dieci posti, pena l'incendio del locale. Siccome l'incentivo a obbedire che ciascuno ha è lo stesso, che gli altri obbediscano o meno (a meno che la disobbedienza di massa distrugga la volontà o il potere di punire esercitato da chi deve far osservare la regola), una regola di questo genere non è una convenzione.

Potrebbe tuttavia essere convenzionale nel senso che forma il contenuto di una convenzione che vale tra coloro che creano le regole. Supponiamo che il comandante di un campo di concentramento che sia eccezionalmente rigido o eccezionalmente indulgente si cacci nei pasticci. Tutti i comandanti hanno allora interesse a non modificare il repertorio di punizioni. Se ciò avviene, il fatto che la punizione per l'assembramento in gruppi di più di sei persone sia di dieci giorni a pane e acqua potrebbe essere una convenzione tra i comandanti. Le regole che prescrivono quella punizione in campi diversi sono, in un certo senso, convenzionali; ma non sono convenzioni.

Altre regole codificano regolarità appartenenti alle specie discusse nel paragrafo 2 come contratti sociali che non sono convenzioni. Se adottiamo l'accezione ristretta di preferenza, in base alla quale le preferenze di una persona non comprendono il suo assenso a obblighi morali, per esempio obblighi di tacito consenso o di comportamento corretto, questa classe di regole risulta ampia. Tali regole prescrivono a ciascun agente un comportamento che può essere in contrasto con le sue preferenze (nell'accezione ristretta),

---

<sup>11</sup>Molto probabilmente un preavviso. La distinzione è di Schelling cfr. *Strategy of Conflict*, pp. 123-124. Un preavviso è un'asserzione secondo la quale, se tu fai A, che a me non piace, con ciò mi darai probabilmente un buon motivo di fare B, che a te non piace. Una minaccia è una dichiarazione della mia attuale intenzione di fare B se tu fai A che io poi abbia o meno un buon motivo di fare B.

ma che è rispondente alle preferenze di tutti gli altri agenti interessati. Sono imperativi ipotetici che stabiliscono quello che uno deve fare per mantenere una promessa tacita, per ricambiare un beneficio, e così via. Possono anche contenere minacce o preavvisi specificanti le sanzioni (istituzionalizzate o informali) alle quali va incontro il trasgressore. La legge penale può consistere in gran parte di regole di questo tipo, perlomeno in una società tradizionale priva di legislazione. Lo stesso vale per parte della nostra morale sociale, per esempio per la regola secondo la quale le promesse vanno mantenute. Lo stesso vale per i regolamenti di una biblioteca relativi alla restituzione dei libri presi in prestito.

Infine, ci possono essere regole che non sono convenzioni soltanto a causa del fatto che vengono imposte per mezzo di sanzioni tanto pesanti da costituire un motivo decisivo di obbedienza anche in caso di disobbedienza da parte di altri. Prendiamo la convenzione con cui i logici potrebbero stabilire una notazione standard; e supponiamo che, dopo che la convenzione vige da vent'anni, qualunque editore respinga un manoscritto che non usi la notazione standard. In questo caso gli editori avrebbero sancito una regola che richiede la notazione standard, una regola imposta con la sanzione della non pubblicazione. Ma non si tratta più di una convenzione, in quanto ogni logico ha una ragione determinante per far uso della notazione standard, che i suoi colleghi la usino o meno. Egli vuole ancora far uso della notazione che usano i suoi colleghi, qualunque essa sia, e sarebbe contento di imitarli se essi la cambiassero tutti; ma non è probabile che si curi di ciò al punto da rinunciare alla pubblicazione. (Presuppongo qui una certa inerzia da parte degli editori, tale che essi potrebbero per un certo tempo insistere con la vecchia notazione standard, anche dopo che la maggior parte dei logici fossero passati a una nuova notazione. Se anche gli editori cambiassero immediatamente, la loro imposizione della notazione standard non toglierebbe nulla alla sua convenzionalità.) In queste condizioni, non è vero che ciascuno preferirebbe far uso di una qualunque notazione non standard se ne facessero uso gli altri; perciò l'uso della notazione standard non è più una convenzione.

Forse ci sono altri tipi di cosiddette regole che non sono convenzioni, ma con questo ho concluso il mio elenco.

Più difficile è sostenere che alcune convenzioni vengono chiamate regole in modo innaturale. (In realtà, è difficile mostrare che c'è una qualunque regolarità che non può essere chiamata regola in qualche contesto.) Ma consideriamo delle convenzioni che coesistono e si trovano in contrasto con regole che sono regole per eccellenza, come in un gioco. Il gioco *Jotto* è definibile come una attività che si conforma alle regole seguenti.

Ci sono due giocatori.

Ciascuno sceglie una parola di cinque lettere da far indovinare all'altro.

Ciascuno propone a turno una parola di prova di cinque lettere di sua scelta, e l'altro gli dice quante lettere la sua parola ha in comune con quella che deve indovinare.

Quando un giocatore propone la parola scelta dall'altro perché la indovini, vince.

Questi sono esemplari perfetti di regole. Definiscono il gioco *Jotto*. Sono facilmente codificate, come sopra, e le loro codificazioni sono adoperate per insegnare a giocare a *Jotto*. La loro violazione sarebbe considerata una prova di incapacità o di mancanza di volontà di giocare a *Jotto*. Sono convenzionali; ma non sono le uniche convenzioni del gioco. Qualunque gruppo di giocatori elaborerà delle intese - convenzioni tacite, spazialmente circoscritte, temporanee, informali - per dare una risposta ai problemi lasciati insoluti dalle regole elencate. Quali sono ammesse tra le parole straniere, le parole dialettali, i nomi propri, le sigle o i neologismi? Può un giocatore ottenere la ripetizione di una risposta precedente (senza perdere il turno) se pensa che ci sia stato un errore? E così via. Volendo, potremmo chiamarle regole di intesa - regole non scritte, regole informali -; ma saremmo anche inclini a sottolineare le loro diversità dalle regole elencate dicendo che non sono regole ma solo convenzioni.

Spero che i miei esempi abbiano lasciato l'impressione che la classe delle cosiddette regole è ibrida, con molti membri dubbi. Potremmo esser tentati di distinguere vari sensi della parola "regola", nella speranza che uno di essi si accordi alla mia definizione di convenzione. Dubito che questo progetto avrebbe successo. Si potrebbero proporre molti sensi, ma probabilmente si rivelerebbero non abbastanza distinti da meritare il nome di sensi diversi. Sembra che abbiamo a che fare con un concetto complesso particolarmente confuso, che varia in relazione al problema in questione, in relazione alle opposizioni che si intende istituire, e in relazione ai pregiudizi filosofici di ciascuno. (A questo punto dovrebbe essere chiaro il motivo per cui ho contrapposto alle convenzioni non le regole ma le "cosiddette regole". Volevo ammettere la possibilità di oscillazione in quello che chiameremmo regola senza specificare se tale oscillazione è ambiguità, vaghezza di confini semantici o altro.)

William Alston parlando, credo, a nome di molti filosofi ha fatto la seguente proposta.

"Al pari della teoria del contratto sociale nella scienza politica,

l'idea secondo cui le parole acquistano per convenzione il loro significato è un mito, se la si prende alla lettera. Ma, al pari della teoria del contratto sociale, essa può costituire la formulazione, in forma mitica, di importanti verità che potrebbero venire formulate in termini più asciutti. È nostra convinzione che queste verità possano venire formulate nella maniera migliore mediante la nozione di regola. Quanto realmente contraddistingue i simboli è dunque il fatto che essi hanno il significato loro proprio in virtù di regole che governano, per ciascuno di loro, e nella comunità di riferimento, il loro uso [...]. D'ora in poi ci sentiremo autorizzati a fare uso del termine 'convenzionale', opportunamente liberato da ogni associazione ingannevole, come abbreviazione di 'sulla base di regole'.<sup>12</sup>

Ma se la mia analisi della convenzione è valida, e se la classe delle cosiddette regole è ibrida come i miei esempi sembrano mostrare, allora è meglio fare esattamente l'opposto: interpretare come convenzioni tacite le "regole di linguaggio" che incontriamo nel nostro lavoro di filosofi del linguaggio. Non abbiamo alcuna giustificazione se ci lasciamo ingannare dalla fuorviante associazione tra convenzione e accordo esplicito.

Né dobbiamo lasciarci ingannare dalle connotazioni fuorvianti della parola "regola". Quando si dice che il linguaggio è governato da regole, noi pensiamo probabilmente a regole che sono state codificate da qualche autorità, o che potrebbero, facilmente venir codificate; a regole che sono imposte per mezzo di sanzioni, formali o informali; a regole che vengono menzionate insegnando o criticando l'uso del linguaggio. Paul Ziff, per esempio, si è lasciato erroneamente indurre allo scetticismo proprio da queste connotazioni fuorvianti, come possiamo vedere dalla sua condanna delle regole del linguaggio:

Io sono interessato alle regolarità: non mi interessano le regole. Le regole non hanno virtualmente nulla a che vedere con la produzione o la comprensione di una lingua naturale.

I filosofi sono portati a raffigurarsi il linguaggio nel modo seguente. Parlare una lingua significa impegnarsi in una attività, un'attività che si accorda a certe regole. Se le regole della lingua vengono violate (o trasgredite, o infrante) lo scopo del linguaggio, la comunicazione, non può essere conseguito se non *per accidens*.

---

<sup>12</sup> *Philosophy of Language*, Prentice-Hall, Englewood Cliffs, 1964, pp. 57-58. [Trad. it. cit., p. 96. (N.d.T.)]

Le regole vengono formulate quando si insegna la lingua, e a esse si fa appello quando si critica l'esecuzione linguistica di qualcuno. La rappresentazione ammette delle variazioni, delle elaborazioni, ma non mi addentrerò maggiormente in questi misteri. Tale rappresentazione del linguaggio non può produrre che confusione, e non può che essere il prodotto di una confusione. L'appello alle regole, mentre si discute quali regolarità vanno riscontrate nel linguaggio, è tanto poco pertinente quanto lo è l'appello alle leggi del Massachusetts mentre si discutono le leggi del moto.<sup>13</sup>

Ma, se consideriamo le regole del linguaggio di cui parlano i filosofi come tacite convenzioni di linguaggio, sfuggiamo all'attacco di Ziff. Noi infatti non supponiamo che queste cosiddette regole siano formulate quando si insegna una lingua - non è neppure necessario supporre che quella lingua venga insegnata -, né supponiamo che a queste cosiddette regole si faccia appello quando si critica l'esecuzione linguistica. Inoltre le nostre regole non saranno soltanto regolarità del comportamento verbale; saranno regolarità del comportamento verbale e delle aspettative e preferenze circa il comportamento verbale, e delle aspettative circa queste aspettative e preferenze, e così via.

### 3.5 Comportamento conformativo

David Shwayder, in *The Stratification of Behaviour*,<sup>14</sup> affronta l'analisi di un concetto di regola. A questo fine introduce un concetto correlativo: quello di *comportamento conformativo*. Il termine è definito esplicitamente senza menzionare le regole. La nozione di regola esposta da Shwayder (che egli spera si identifichi con la nostra nozione ordinaria) è quindi definita esplicitamente in termini di comportamento conformativo. I due concetti vanno messi in relazione nel modo seguente: le regole sono certi fatti che possono essere motivi del comportamento di un agente; "il comportamento conformativo è di una specie tale che richiede che l'agente abbia un certo tipo di motivo o di falso motivo, che può essere schematizzato con le parole 'Questa è la regola' (p. 238). Questa definizione include anche, volutamente, il *non* conformarsi a una regola supposta; il comportamento *conformativo* non si limita al comportamento *che si conforma*.

La teoria di Shwayder relativa alla regola e al comportamento conformativo assomiglia molto alla mia analisi della convenzione. (Shwayder usa

---

<sup>13</sup> *Semantic Analysis*, Cornell University Press, Ithaca, New York, 1960, pp. 34-35.

<sup>14</sup> Humanities Press, New York, 1965.

anche il termine “convenzione”, ma per indicare qualcosa di completamente diverso. Dato che la sua “convenzione” è molto simile al “significato<sub>nn</sub>” di Grice, che discuterò nel capitolo 4.5, non mi occuperò più di essa.) Il comportamento che, in base alla mia definizione, è qualificabile come conformità a una convenzione, in base alla definizione di Shwayder si qualifica come comportamento conformativo, in gran parte per gli stessi motivi. Ma emergono alcune differenze. Nel comportamento conformativo rientra la conformità a alcune regolarità che non sono convenzioni e che a malapena potrebbero esser dette regole. Shwayder arriva gradatamente alla sua definizione di regola e di comportamento conformativo formulando diverse tesi relative alle regole che dovrebbero derivare da qualunque teoria soddisfacente, e che deriveranno dalla sua.

- (1) Il comportamento conformativo non è semplicemente un comportamento che si conforma a regole. “L’agente stesso deve o conformarsi o violare la regola. Una condizione perché ciò avvenga è che egli creda che una regola esiste” (p. 241).
- (2) Alcune regole devono essere formulate anticipatamente; altre no. Ma, in ogni caso, la regola è distinta dalla formulazione o enunciazione di quella regola.
- (3) Le regole non sono, o non sono soltanto, regolarità di comportamento. Ancor meno sono generalizzazioni che possiamo formulare riguardo alle regolarità di comportamento. Ma le regolarità di comportamento possono essere dovute all’esistenza di regole alle quali gli agenti si conformano regolarmente.
- (4) Le regole sono innanzitutto regole di una comunità di agenti. Il comportamento conformativo presuppone regole comunitarie. Le regole private sono possibili, ma sono in qualche modo secondarie rispetto alle regole comunitarie, ed è essenziale che siano passibili di diventare regole comunitarie.
- (5) Le regole sono motivi certi di azione: sono fatti di una certa specie la cui supposta esistenza è motivo di un certo comportamento, cioè del comportamento conformativo. (È così anche nei casi atipici in cui l’agente si sbaglia riguardo all’esistenza di una regola, o in cui viola una regola.) Una regola è un motivo d’azione in virtù di alcuni principi del tipo “ci si deve comportare in conformità alle regole”, “ci si deve comportare in modo da evitare le punizioni o i disagi che possono derivare dalla violazione o dall’infrazione di una regola”, e così via (p. 251).

Giungiamo così alla tesi centrale di Shwayder relativa alla natura delle regole: “Le regole comunitarie sono sistemi di aspettative. Un agente si conforma a una regola di questo tipo se agisce per il motivo che i membri della comunità sono autorizzati ad aspettarsi che egli si comporti cosa” (p. 252). Egli sviluppa cosa la sua idea:

“Limitandoci alle regole comunitarie, l’idea è questa: una persona segue una regola se si conforma a quelle che capisce essere le aspettative degli altri; e l’esistenza di una regola è, inoltre, ciò che autorizza le aspettative degli altri, rendendole in tal modo “legittime”. Una regola comunitaria esiste se i membri di una comunità regolano la propria condotta in base a quello che gli altri membri della comunità si aspetterebbero legittimamente da loro. La regola è costituita dalle aspettative alle quali ci si conforma e al tempo stesso da ciò che legittima o giustifica quelle aspettative. La regola è, per così dire, un sistema di aspettative reciproche della comunità. Quando ci si conforma a una regola, si agisce sapendo o credendo che gli altri si aspettino che noi ci comportiamo così. Il fatto che gli altri siano autorizzati ad avere quelle aspettative costituisce la nostra motivazione. Naturalmente uno può comportarsi in contrasto con tali regole; ma perfino in questo caso deve credere che gli altri abbiano delle aspettative legittime. Se uno non ha alcuna idea riguardo a quello che ci si aspetta da lui, egli non può agire né in conformità né in contrasto con la regola” (p. 253).

Ma Shwayder non è ancora soddisfatto del modo in cui ha definito il comportamento conformativo, in quanto menziona ancora le regole. Deve trovare il modo di riformulare la condizione, inaccettabile, secondo la quale “l’esistenza di una regola è... ciò che autorizza le aspettative degli altri, rendendole in tal modo ‘legittime’”. Si tratterebbe cioè dell’esistenza di qualche aspetto della regola: un aspetto che possa essere descritto in altri termini.

Tuttavia, prima di proseguire, Shwayder si limita provvisoriamente a “quello che è il caso sicuramente più fondamentale e costantemente più frequente, cioè il caso delle regole comunitarie, con le condizioni addizionali (1) che alcuni dei membri di quella comunità siano presenti, laddove (2) gli agenti agiscono in modo da conformarsi, e (3) che si possa assumere che l’agente e i suoi osservatori sappiano tutto quello che importa sapere della situazione” (p. 254). (Gli “osservatori” sono estranei che descrivono il comportamento di cui Shwayder sta esaminando il concetto. Così la (3) non è

una condizione di conoscenza comune, e Shwayder è soggetto alle obiezioni che ci hanno indotto a stabilire una condizione di questo genere.

In effetti, il comportamento conformativo, nel caso fondamentale, è stato finora descritto come un comportamento giustificabile per mezzo del ragionamento rappresentato nella Figura 31. (Dove Shwayder parla di conoscenza,



io continuo a parlare di aspettative. Ma la differenza non è importante, poiché in qualunque caso normale di aspettativa ci sarà anche conoscenza.) Il problema è ridescrivere questo ragionamento in un modo che non faccia menzione della regola.

La soluzione di Shwayder è contenuta nella sua formula: “Io agisco in base alla conoscenza del fatto che gli altri sanno che io mi comporterò in base alla conoscenza del fatto che essi si aspettano che io mi comporti così” (p. 256). Penso che ciò significhi che parte delle giustificazioni dell’agente devono essere rappresentate come nella Figura 32. Combinando questi due



Figura 32

frammenti della giustificazione dell’agente scopriamo che la sua giustificazione deve essere rappresentata in parte dalla Figura 33. Penso che un’azione soddisfi la definizione di comportamento conformativo data da Shwayder se e solo se può essere giustificata da un ragionamento che segua questo schema, che chiamerò *schema di comportamento conformativo*.

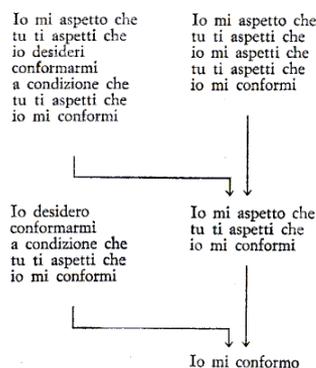


Figura 33

Shwayder ha introdotto le seguenti modifiche nella condizione da lui imposta al comportamento conformativo. Nella prima versione, io so che le aspettative degli altri riguardo al mio comportamento sono giustificate dall'“esistenza di una regola”. Nella seconda versione, io so che l'aspettativa degli altri riguardo al mio comportamento è giustificata dal fatto che essi potrebbero acquisirla derivandola dalla loro conoscenza del fatto che io so che essi si aspettano che io mi comporti così, unita alla loro conoscenza del fatto che io cercherò di fare ciò che essi si attendono da me. Inoltre, è con la ricostruzione proprio di questa derivazione che *io* potrei acquisire la conoscenza delle loro aspettative circa il mio comportamento.

Con questa modifica Shwayder è riuscito a eliminare, analizzandola, anche l'ultima menzione della regola. Adesso che la sua definizione del comportamento conformativo nel caso fondamentale è soddisfacente, egli è libero di passare a definire, senza circolarità, le *regole* come quelle aspettative reciproche circa il comportamento che figurano nel comportamento conformativo in quanto motivi di azione. Esse sono esattamente quelle aspettative che gli altri membri della comunità dell'agente hanno derivato nel modo specificato.

Shwayder estende poi le sue definizioni indebolendo le restrizioni imposte al caso fondamentale. Egli definisce un comportamento conformativo in cui nessun altro membro della comunità dell'agente è presente; in cui l'agente viola una regola; in cui l'agente si sbaglia riguardo all'esistenza di una regola; o in cui la regola in questione è una regola privata, potenzialmente pubblica. Ma noi non seguiremo queste estensioni. Infatti è la teoria di Shwayder sul comportamento conformativo nel caso fondamentale che si avvicina di più alla mia analisi della conformità a una convenzione.



Figura 34

Se prendiamo un esempio qualunque di azione che, sulla base della mia analisi, si conforma a una convenzione, scopriremo che soddisfa anche la definizione data da Shwayder di comportamento conformativo nel caso fondamentale. Non è un caso. Ogniquale volta io mi conformo a una convenzione, la mia azione è giustificabile per mezzo di ricostruzioni; la profondità degli incastri è limitata soltanto dalla disponibilità di premesse ausiliari riguardanti la razionalità. Considerando, per semplicità, il caso con due persone, e trascurando l'uso delle premesse di razionalità, la mia giustificazione può essere rappresentata per mezzo del solito schema di ricostruzione, come illustrato nella Figura 34. Ma c'è un modo diverso di rappresentare sostanzialmente la stessa giustificazione della mia azione. Consideriamo due stadi consecutivi qualsiasi del nostro schema. Presi insieme, essi mi fanno passa-

re da un'aspettativa circa l'azione di ordine  $(n + 2)$ -esimo a un'aspettativa circa l'azione di ordine  $n$ -esimo (o, se  $n = 0$ , circa una decisione di agire), con un'aspettativa di ordine  $(n + 1)$ -esimo come intermediaria. Ma lo stesso ragionamento (stesse premesse, stessa conclusione) potrebbe essere sviluppato seguendo un ordine diverso, con un diverso passaggio intermedio. Gli ultimi due stadi, per esempio, potrebbero venir ordinati come nella Figura 35. Le stesse premesse mi inducono alla stessa azione, ma le premesse

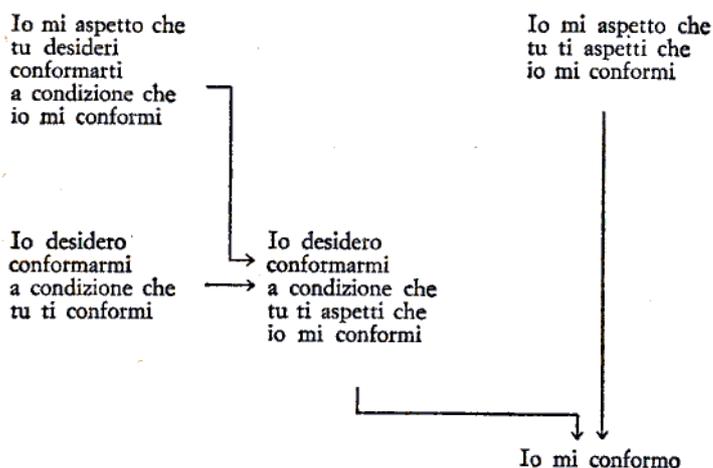


Figura 35

vengono usate in un ordine diverso. C'è un nuovo passaggio intermedio: io derivo il desiderio di conformarmi a condizione che tu ti aspetti che io lo faccia. Dato questo, la mia aspettativa che tu ti conformi diventa superflua. Riordinando ogni coppia di stadi a partire dal basso verso l'alto, otteniamo lo schema di ricostruzione riordinato della Figura 36. E a questo punto possiamo renderci conto di come lo schema del comportamento conformativo sia proprio la parte centrale (dentro la linea tratteggiata) dello schema di ricostruzione riordinato. Ne segue che qualunque azione conforme a una convenzione, essendo giustificabile per mezzo di ricostruzioni e quindi di ricostruzioni riordinate, soddisfa la definizione di comportamento conformativo data da Shwayder.

Non vale, però, la converso. Supponiamo che io voglia conformarmi se gli

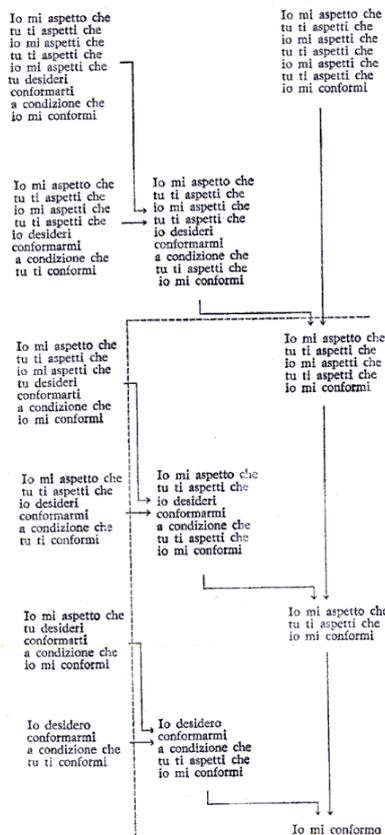


Figura 36

altri si aspettano che mi conformi, ma indipendentemente da qualunque cosa che essi faranno per il fatto che si aspettano che io mi conformi. In questo caso la giustificazione della mia azione non può adattarsi allo schema di ricostruzione riordinato. Siccome la colonna di sinistra sarà mancante, la mia azione non può essere conforme a una convenzione. Ma la mia giustificazione potrebbe ancora adattarsi allo schema di comportamento conformativo. Se ciò avvenisse, la mia azione sarebbe un comportamento conformativo. Io sarei l'unico *agente* della situazione; gli altri sarebbero implicati solo in quanto supposti detentori di aspettative nei miei riguardi. Potrebbe darsi che io stia soltanto cercando di dimostrare a me stesso che, se voglio, posso vivere in accordo con le aspettative degli altri. Potrebbe darsi che io desideri

farmi ritenere da loro un tipo noioso e prevedibile. Potrebbe darsi che voglia evitare di deluderli o di sorprenderli. In ogni caso, potrebbe darsi che mi comporti così “in base alla conoscenza del fatto che gli altri sanno che io mi comporterò in base alla conoscenza del fatto che essi si aspettano che io mi comporti così”. Ma non sembra che siano questi i casi di comportamento conformativo che Shwayder aveva in mente. Essi non si adattano alla sua formula “Agisco per il motivo che ‘Questa è la regola’”. In realtà agisco per il motivo che “Questa è la loro aspettativa legittima” (e la legittimità è della specie giusta), ma in tal caso la loro aspettativa legittima non si può eventualmente chiamare regola.

Richiedendo soltanto un frammento centrale, e non la totalità, dello schema di ricostruzione riordinato, Shwayder ha ottenuto una certa generalità. Ma non è chiaro perché questa generalità sia desiderabile. E, per raggiungerla, ha trascurato un fatto importante relativo alla specie di comportamento conformativo in questione: il fatto che io voglio conformarmi alle tue aspettative a causa del modo in cui mi aspetto che tu agisca sulla base delle tue aspettative. Se io pensassi che tu non agiresti sulla base delle tue aspettative, mi interesserei del modo in cui agiresti, non di quello che ti aspetti. Se questo fatto viene trascurato, la nostra comprensione del fenomeno risulta gravemente distorta.

C'è un secondo modo in cui un'azione può essere un comportamento conformativo senza essere conforme a una convenzione, anche se la giustificazione dell'azione si adatta a tutto lo schema di ricostruzione riordinato. La giustificazione per mezzo di ricostruzioni, e quindi la giustificazione per mezzo di ricostruzioni riordinate, si applica a qualunque azione che sia parte di un equilibrio proprio. Ma l'equilibrio non deve necessariamente essere in un problema di coordinazione, e non deve necessariamente essere un equilibrio di coordinazione. Può essere perfino un equilibrio unico di un gioco di conflitto puro.

Supponiamo che, per qualche ragione, dobbiamo giocare spesso, a coppie, a testa o croce, con la possibilità di “passare”. È un gioco di conflitto puro illustrato dalla matrice delle vincite della Figura 37: il giocatore di riga e quello di colonna devono giocare simultaneamente passando (R1 o C1), mettendo una moneta in posizione di “testa” (R2 o C2), o mettendo una in posizione di “croce” (R3 o C3). Se le posizioni delle monete concordano, il giocatore di colonna le vince tutte e due; se non concordano, le vince il giocatore di riga; se entrambi i giocatori passano, ciascuno si tiene la sua moneta; se uno solo passa, l'altro gli paga una moneta che vale la metà di quella in gioco. C'è un equilibrio unico  $\langle R1, C1 \rangle$ ; non è un equilibrio di coordinazione.

	C1	C2	C3
R1	0	-0,5	-0,5
R2	0,5	1	-1
R3	-0,5	-1	1
	0	0,5	0,5
	-0,5	-1	1
	-0,5	1	-1

Figura 37

Io passo, a meno che mi senta sicuro che tu giocherai testa (o croce), nel qual caso giocherò croce o testa (secondo i casi) per batterti. In particolare, preferisco passare se tu passi. Se giochiamo spesso, allora, una volta che la natura del gioco è diventata per noi conoscenza comune, passeremo ogni volta. Tra noi è conoscenza comune che passiamo, e che lo facciamo perché ci aspettiamo reciprocamente di farlo.

Quando io passo, la mia azione è un comportamento conformativo. La sua giustificazione si adatta allo schema di ricostruzione riordinato e *a fortiori* allo schema di comportamento conformativo. Io agisco sulla base della conoscenza del fatto che tu sei autorizzato ad aspettarti che io passi, o - più precisamente - del fatto che tu sai che io agirò sulla base della conoscenza del fatto che tu ti aspetti che io passi.

Ma la nostra regolarità nel passare non soddisfa la definizione di convenzione. La sua esclusione è giustificata da due importanti differenze tra essa e i casi tipici di convenzione: (1) Gli equilibri dei giocatori non sono equilibri di coordinazione. Non rappresentano una cooperazione nell'interesse comune, bensì compromessi statici tra interessi completamente opposti. Nessuno dei due è soddisfatto da  $\langle R1, C1 \rangle$ . Nessuno dei due avrebbe potuto conseguire un risultato migliore se si fosse comportato diversamente, ma ciascuno desiderava che l'altro si comportasse diversamente. (2) Gli equilibri dei giocatori sono gli unici possibili. In essi non c'è nulla di arbitrario. Non esiste un altro modo possibile di giocare con una strategia valida e con precise attese reciproche. Né sarebbe naturale chiamare regola la nostra regolarità, se non nel senso in cui qualunque regolarità o qualunque massima strategica può a volte essere chiamata regola (e Shwayder non intende certo usare

un'accezione così comprensiva). Deve trattarsi di un'altra specie di comportamento conformativo trascurato da Shwayder, in quanto - intuitivamente - non è un'azione compiuta per il motivo che "Questa è la regola".

Un'ultima differenza tra le regole di Shwayder e le mie convenzioni merita di essere menzionata solo perché può darsi che il lettore l'abbia già notata. Per Shwayder una regola è un sistema di aspettative atte a produrre un comportamento regolare. Per me la convenzione è una regolarità di comportamento prodotta da un sistema di aspettative. Shwayder intende permettere (in una delle estensioni operate a partire dal caso fondamentale) la frequente infrazione delle regole; questa generalità può essere plausibile per il suo *analysandum* "regola", ma non per il mio *analysandum* "convenzione". Io sono interessato ai casi in cui abbiamo sia la regolarità che le aspettative, e in tali casi la regolarità è la cosa centrale. Ma la differenza è superficiale. Svanirebbe se riformulassimo le nostre analisi nel modo seguente: "Una regola (convenzione) esiste se e solo se..."; non "Una regola (convenzione) è..."

### 3.6 Imitazione

Chi è partecipe di una convenzione si conforma a una regolarità perché ha interesse a conformarsi se certi altri lo fanno e perché crede, a ragione, che essi lo facciano. Agisce come agisce perché si aspetta che gli altri agiscano così. In breve, li imita. Ma non concluderemmo per questo che qualunque regolarità che inizia o persiste in virtù di qualche specie di imitazione reciproca è una convenzione. Ci sono parecchi casi in cui ogni membro di una popolazione si comporta conformemente a una regolarità perché lo fanno gli altri, senza che per questo tale regolarità sia una convenzione.

A volte gli uomini si copiano reciprocamente - per esempio nei tic - più o meno senza sapere che lo fanno. Dato un gruppo composto esclusivamente di persone di questo genere, un tic può diffondersi e durare per imitazione reciproca. Ma non c'è preferenza da parte di alcuno (a meno che consideriamo ogni inclinazione come una preferenza transitoria). Ciascuno fa semplicemente qualcosa, senza curarsene e quasi inconsapevolmente. Così, *a fortiori*, comportandosi in tal modo lo fanno gli altri, la sua azione non risponde a alcun interesse. La regolarità prodotta non è una convenzione.

A volte gli uomini copiano le loro reciproche preferenze. In mezzo a bevitori di tè, un bevitore di caffè può in qualche modo mettersi a preferire il tè. Non è che egli preferisca bere caffè se gli altri bevono caffè e bere tè se gli altri bevono tè. Dapprima preferiva il caffè, indipendentemente da quello

che preferivano gli altri; in seguito preferisce il tè altrettanto incondizionatamente. Ma è stata la sua esposizione ai bevitori di tè che l'ha fatto cambiare. Osservando il processo di adattamento di un detenuto, possiamo dire di avere a che fare con preferenze non condizionali mutevoli, non con preferenze condizionali fisse; lo stesso vale quando osserviamo le sue preferenze relative a qualcosa come una scommessa che gli garantisca un rifornimento illimitato di tè se tutti i suoi vicini passano al caffè. Dato un gruppo composto esclusivamente di copiatori di preferenze, le preferenze e le azioni a esse rispondenti potrebbero diffondersi e durare per imitazione reciproca. Ma la regolarità così prodotta non sarebbe una convenzione, in quanto le azioni conformi a essa risponderebbero a una preferenza non condizionale.

A volte gli uomini si fidano reciprocamente dei loro giudizi pratici: attribuendo agli altri dei probabili buoni motivi per quello che stanno facendo, io posso dedurre dalle loro azioni che essi sanno qualcosa che io non so: qualcosa che costituisce un buon motivo perché io faccia lo stesso. Posso indossare l'impermeabile perché lo fanno gli altri, pensando che probabilmente indossano l'impermeabile perché hanno udito una previsione di pioggia. Prima e dopo aver visto i miei vicini in impermeabile, la mia preferenza era di indossare l'impermeabile se e solo se stava per piovere. E questa è la mia preferenza indipendentemente da quello che fanno gli altri. Il fatto che essi indossino l'impermeabile mi induce a indossare il mio perché considero le loro azioni come una prova: la prova del fatto che si aspettano la pioggia e quindi (indirettamente, attraverso le mie attuali credenze relative alle cause probabili delle loro aspettative), la prova della probabilità della pioggia. Una regolarità potrebbe diffondersi e durare in un gruppo di persone proprio in virtù di questo tipo di imitazione reciproca. Un giorno potrebbe accadere che *ciascuno* indossi l'impermeabile perché vede che gli altri lo indossano e ne deduce - abbastanza ragionevolmente, forse,<sup>15</sup> ma erroneamente - che essi hanno probabilmente udito una previsione di pioggia. (Colui che l'ha indossato per primo deve essere, dapprima, un'eccezione, in quanto ha indossato l'impermeabile quando gli altri non lo indossavano ancora. Ma, qualunque fosse la sua ragione originaria, se continua a indossare l'impermeabile perché vede che lo indossano gli altri, il suo comportamento diventa esattamente uguale a quello degli altri.) Questa regolarità non è una convenzione; la preferenza su cui si fonda non è subordinata alla conformità degli altri.

---

<sup>15</sup>Può darsi che ciascuno fosse perfettamente ragionevole nel trarre inferenze e nel comportarsi come ha fatto - anche se nessuno lo penserà quando apprenderà quello che è accaduto. L'evidente irrazionalità del gruppo può non esser dovuta a irrazionalità dei suoi membri. Non è un errore aspettarsi la pioggia quando si vede la gente con l'impermeabile, nonostante i cattivi risultati prodotti questa volta da tale comportamento.

Anche quando delle persone si imitano reciprocamente a causa della loro preferenza condizionale a fare qualcosa se lo fanno gli altri, la regolarità che persiste in virtù di questa imitazione non è ancora necessariamente una convenzione. Può darsi, infatti, che non si tratti di una situazione in cui la coincidenza di interessi predomina sul conflitto. Viene mantenuta una qualche specie di equilibrio, ma non si tratta di un equilibrio di *coordinazione*. Per esempio, nel caso di conflitto puro del paragrafo precedente, ciascuno ha la preferenza condizionale a passare se l'altro passa. Questa preferenza rende possibile una regolarità in base alla quale ciascuno passa tutte le volte perché passa l'altro. Ma abbiamo visto che questa regolarità non è una convenzione; sebbene ciascuno sia soddisfatto della propria scelta (data quella dell'avversario), nessuno dei due è soddisfatto della scelta dell'avversario (data la propria scelta). Ciascuno vuole conformarsi alla regolarità consistente nel passare sempre se anche l'altro si conforma a essa; ma ciascuno preferirebbe conformarsi anche nel caso che l'altro non si conformi. Qualora si tratti invece di una convenzione genuina, ciascuno vuole conformarsi se lo fanno gli altri, e ciascuno vuole che gli altri si conformino se si conforma lui.

Abbiamo così distinto cinque specie pure di regolarità basata sull'imitazione reciproca: la convenzione stessa e quattro contraffazioni. Esse differiscono per il tipo di imitazione che entra in gioco, ossia per il processo grazie al quale certe azioni producono azioni simili da parte di altri. Ciascuno dei casi considerati finora presenta un unico tipo di imitazione; ma potrebbero benissimo esserci degli ibridi, delle regolarità basate su imitazione reciproca di parecchi tipi diversi mescolati in proporzioni variabili. Sembra plausibile che le usanze, le mode, le psicosi collettive, le sommosse e la popolarità, quali li conosciamo, siano il prodotto dell'azione congiunta di numerosi tipi di imitazione reciproca. Supponiamo di indossare, quest'anno cravatte nocciola perché le indossano gli altri. Morgan segue la moda senza essere cosciente di seguirla; quando sceglie casualmente una cravatta, gli capita di sceglierne proprio una nocciola. Jones indossa cravatte nocciola perché gli piace il colore; ma, a sua insaputa, i suoi gusti sono determinati dalla moda prevalente e cambieranno con essa. Griffith indossa la cravatta nocciola perché suppone erroneamente che le altre persone che si vestono di nocciola abbiano scoperto qualche particolare potere posseduto da questo colore, potere del quale beneficerà anch'egli. Owen e Thomas desiderano indossare entrambi qualunque colore di cravatta adottato dagli altri; ma Owen spera di trovare degli anticonformisti occasionali di cui ridere, mentre Thomas spera che non ci sia alcun anticonformista. Ciascuno segue a suo modo la moda, e l'uso delle cravatte nocciola persiste e si diffonde in virtù di tipi diversi di imitazione reciproca.

Tra i casi di regolarità per imitazione reciproca non annovero il fatto che un'idea geniale diventi popolare. Tale regolarità si diffonde per imitazione, via via che la gente vede quello che fanno gli altri e capisce che, comportandosi in modo analogo, ne trarrebbe dei benefici. Ma l'imitazione non è reciproca - due persone non imparano il trucco l'una dall'altra - e la regolarità non persiste in virtù dell'imitazione. Ciascuno, una volta che ha cominciato, continua perché ricava dei vantaggi da quello che fa, indipendentemente dal fatto che gli altri continuino o meno.

## Capitolo 4

# Convenzione e comunicazione

### 4.1 Segnali campione

Finora abbiamo considerato le convenzioni in generale. Adesso passeremo a una classe particolarmente importante di convenzioni: quelle per mezzo delle quali assegniamo a un'azione opportuna la prerogativa di essere un segnale.

La comunicazione mediante segnali convenzionali è un fenomeno banale, tanto che dobbiamo fare uno sforzo per non darlo per scontato. Potremmo esercitare a piacere la nostra tacita comprensione di esso senza mai renderla più esplicita. Questo è ciò che accadrebbe se cominciassimo dicendo che le azioni sono segnali quando le forniamo di significati. Tale ovvietà non ci farà progredire nel tentativo di descrivere il fenomeno della segnalazione indipendentemente dalla nostra precedente comprensione tacita di esso. Descriveremo quindi il fenomeno in altri termini, e lasceremo che il significato cerchi se stesso.

Consideriamo un comunicante e il suo uditorio: per esempio, il sagrestano della Old North Church e Paul Revere. Il sagrestano agisce in base a un qualche piano circostanziale del tipo:

R1: Se si scorge che i soldati inglesi non si spostano, non appendere lanterne nella cella campanaria. Se si scorge che i soldati inglesi si spostano via terra, appendere una lanterna nella cella campanaria. Se si scorge che i soldati inglesi si spostano via mare, appendere due lanterne nella cella campanaria.

Oppure:

R2: Se si scorge che i soldati inglesi non si spostano, appendere una lanterna nella cella campanaria. Se si scorge che i soldati inglesi si spostano via terra, appendere due lanterne nella cella campanaria. Se si scorge che i soldati inglesi si spostano via mare, non appendere lanterne nella cella campanaria.

Oppure:

R3: Se si scorge che i soldati inglesi non si spostano, appendere una lanterna nella cella campanaria. Se si scorge che i soldati inglesi si spostano via terra, non appendere lanterne nella cella campanaria. Se si scorge che i soldati inglesi si spostano via mare, appendere due lanterne nella cella campanaria.

(Ci sono altri tre piani circostanziali con nessuna lanterna, una lanterna e due lanterne, più un numero qualunque di piani ulteriori che prevedono altre azioni - appendere tre lanterne, appendere lanterne colorate, agitare delle lanterne, spiegare una bandiera, ecc.) Paul Revere agisce in base a un piano circostanziale del tipo:

C1: Se non si scorge alcuna lanterna appesa nella cella campanaria, tornare. Se si scorge una lanterna appesa nella cella campanaria, avvertire la popolazione che i soldati inglesi arrivano via terra. Se si scorgono due lanterne appese nella cella campanaria, avvertire la popolazione che i soldati inglesi arrivano via mare.

Oppure:

C2: Se non si scorge alcuna lanterna appesa nella cella campanaria, avvertire la popolazione che i soldati inglesi si spostano via mare. Se si scorge una lanterna appesa nella cella campanaria, tornare. Se si scorgono due lanterne appese nella cella campanaria, avvisare la popolazione che i soldati inglesi arrivano via terra.

Oppure:

C3: Se non si scorge alcuna lanterna appesa nella cella campanaria, avvisare la popolazione che i soldati inglesi arrivano via terra. Se si scorge una lanterna appesa nella cella campanaria, tornare. Se si scorgono due lanterne appese nella cella campanaria, avvisare la popolazione che i soldati inglesi arrivano via mare.

A ciascuno dei due importa poco quale piano circostanziale seguire, pur-

ché i loro due piani si combinino in modo da garantire che Paul Revere avvisi la popolazione che i soldati inglesi arrivano via terra se e solo se il sagrestano li scorge spostarsi via terra, e che Paul Revere avvisi la popolazione che i soldati inglesi arrivano via mare se e solo se il sagrestano li scorge spostarsi via mare (entro i limiti consentiti dalle interferenze e dagli errori). Ciascuno deve agire in base a un piano circostanziale scelto tenendo conto della propria aspettativa circa la scelta dell'altro: un'aspettativa che si è formata, in questo caso, per accordo esplicito. La loro situazione è un problema di coordinazione, illustrato nella Figura 38. Se essi raggiungono uno degli

	C1	C2	C3	...
R1	1	0	0,5	
R2	0	1	0,5	
R3	0,5	0,5	1	
⋮	0,5	0,5	1	

Figura 38

equilibri di coordinazione (per esempio gli equilibri storici  $\langle R1, C1 \rangle$ , o  $\langle R2, C2 \rangle$ ,  $\langle R3, C3 \rangle$ ...) si ha una comunicazione riuscita. Queste combinazioni garantiscono che Paul Revere darà l'avviso giusto. Se essi raggiungono una delle combinazioni del tipo  $\langle R1, C3 \rangle$ , si verifica una comunicazione riuscita condizionalmente: Paul Revere può dare o non dare l'avviso giusto, a seconda di quello che fanno i soldati inglesi. I fallimenti completi sono le combinazioni del tipo  $\langle R1, C2 \rangle$ , che garantiscono che Paul Revere darà in qualunque caso l'avviso sbagliato.

Ho appena descritto le caratteristiche di un caso di segnalazione senza menzionare il significato dei segnali, ossia il fatto che due lanterne significavano che i soldati inglesi stavano arrivando via mare, o qualunque altra cosa. Ma non sembra che sia stato taciuto alcunché di importante, per cui ciò che è stato detto deve in qualche modo implicare che i segnali hanno i loro propri significati.

Non dovrebbe darci fastidio il fatto che quello in questione sia un problema di coordinazione di piani circostanziali e non - come nei casi precedenti - di coordinazione delle azioni stesse. Stiamo considerando la *scelta* di un piano circostanziale compiuta dall'agente come un'azione: ossia un'azione che è parte di un'altra consistente nello scegliere appunto un piano circostanziale per poi conformarvisi. L'azione di un agente in base a un piano circostanziale è concepita come consistente di due fasi: scelta di un piano circostanziale ed esecuzione del piano prescelto. La scelta dell'agente è concentrata nella prima fase; il suo uso della sua capacità di dire quale circostanza si verifica è concentrato nella seconda fase. Il problema di coordinazione degli agenti riguarda soltanto la prima fase. Per esempio, se il sagrestano appende due lanterne dopo aver scorto i soldati inglesi che si spostano via mare, può darsi che agisca così in base a R1, o a R3 o a un qualunque numero di altri piani circostanziali. Noi assumiamo che la differenza di piano circostanziale provochi una differenza di azione, anche se si tratta di una differenza che si rivela irrilevante.

In questo caso, il sagrestano e Paul Revere si sono accordati sui segnali relativi a ogni singola occasione. In altri casi, lo stesso sistema di segnalazione - consistente di punti di equilibrio di coordinazione di piani circostanziali che sono sostanzialmente analoghi per un comunicante e un uditorio - può presentarsi ripetutamente, senza bisogno di un nuovo accordo ogni volta. La regolarità per mezzo della quale i comunicanti e gli uditori usano una tale coppia di piani circostanziali è una convenzione.

Consideriamo sei esempi di convenzioni di segnalazione.

(1) Il Codice internazionale dei segnali elenca una serie di corrispondenze tra esposizioni di bandiera e certe situazioni. In altri termini, fornisce un piano circostanziale per le navi (precisamente, per gli ufficiali di marina) della forma: in una situazione così-e-così, issare la bandiera così-e-così. C'è un piano circostanziale complementare per le navi che vedono delle bandiere issate sulle navi vicine: se una nave issa una bandiera così-e-così, comportarsi in modo appropriato all'ipotesi che la nave si trovi nella situazione così-e-così. Le navi si comportano regolarmente in base a questi due piani complementari. Una nave che sta compiendo una prova di velocità issa una bandiera bianca a sinistra e blu a destra, e le altre navi, vedendo la bandiera, si tengono lontane dalla sua rotta. Una nave che carica o scarica esplosivi issa una bandiera bianca a sinistra e rossa a destra, e le altre navi, vedendo la bandiera, prendono le precauzioni del caso. Una nave il cui equipaggio si sia ammutinato issa una bandiera con una croce gialla in campo rosso sopra una bandiera a strisce diagonali gialle e rosse, e le altre navi, vedendo quelle

bandiere, portano tutto l'aiuto possibile.

La regolarità consistente nell'agire in accordo a questa coppia di piani circostanziali è una convenzione. Qualunque nave vi si conforma quando si trova in una delle situazioni designate o quando vede delle bandiere sulle navi vicine. E si conforma perché si aspetta che si conformino anche le altre navi, in particolare, quelle altre navi che sono con essa nella relazione del comunicante con l'uditorio o dell'uditorio con il comunicante; e preferisce conformarsi a questa regolarità, o a una simile, se esse si conformano. Esse condividono questa preferenza. E tra le navi è conoscenza comune che queste condizioni valgono.

La convenzione che regola le esposizioni delle bandiere è relativamente formale. Ha avuto origine per accordo esplicito tra i rappresentanti di coloro che ora vi partecipano; è codificata esplicitamente; le violazioni sono punibili indipendentemente dalla conformità da parte degli altri. Per esempio, una nave potrebbe essere ritenuta responsabile di un incidente se non avesse issato la bandiera prescritta, anche se quella parte del codice fosse caduta in disuso. Ma, se è importante che le navi rispondano in modo appropriato alle situazioni delle altre, allora, ancora una volta, questo incentivo indipendente sarebbe superfluo se concordasse con la convenzione, mentre sarebbe superato in efficacia se non concordasse; in ogni caso non sarebbe decisivo, e quindi non toglierebbe nulla alla convenzionalità del codice.

(2) Passiamo adesso a una convenzione di segnalazione molto informale: quella per mezzo della quale un aiutante, stando dietro un autocarro, fa dei gesti al guidatore per aiutarlo a manovrare l'autocarro in uno spazio ristretto. Qui non c'è alcun accordo esplicito, nessun autorevole testo di riferimento, nessuna imposizione istituzionalizzata, nessun insegnamento. Ma ci sono gli elementi di un sistema di segnali convenzionali: i gesti dell'aiutante dipendono dalla posizione dell'autocarro e la manovra del guidatore dipende dai gesti dell'aiutante da lui osservati. Così la manovra dipende indirettamente dalla posizione dell'autocarro, e questa dipendenza è tale da soddisfare l'interesse comune dell'aiutante e del guidatore a far entrare l'autocarro nello spazio ristretto.

L'aiutante avrebbe potuto far sì che i suoi gesti dipendessero dalla posizione dell'autocarro in un modo diverso. Il guidatore avrebbe potuto far sì che la sua manovra dipendesse dai gesti dell'aiutante in un modo diverso. Alcune delle altre combinazioni di dipendenze avrebbero prodotto la stessa dipendenza della manovra dalla posizione dell'autocarro, e avrebbero contribuito in misura pressoché uguale a far entrare l'autocarro nello spazio. L'aiutante gesticola come gesticola perché si aspetta che il guidatore

risponda come risponde, e il guidatore risponde come risponde perché si aspetta che l'aiutante gesticoli come gesticola. Essi hanno derivato queste aspettative dalla loro esperienza passata con aiutanti e guidatori, i quali hanno stabilito un precedente gesticolando e rispondendo in base alle stesse dipendenze.

Il loro comportamento si conforma a una regolarità convenzionale. È una regolarità che sarebbe difficile descrivere, se non altro perché è difficile descrivere un gesto o una manovra in modo sufficiente a identificarli. Eppure gli aiutanti e i guidatori esperti si conformano, si aspettano la conformità e riconoscono la conformità, presumibilmente senza l'aiuto di descrizioni del loro comportamento regolare.

Possiamo dire che gli aiutanti e i guidatori agiscono in base ai loro rispettivi piani circostanziali, a condizione che ci rendiamo conto che questi piani sono semplicemente descrizioni che *si potrebbero dare* del modo in cui le loro azioni dipendono dalle loro osservazioni. Nell'esempio (1) i piani circostanziali del comunicante e dell'uditorio costituivano qualcosa di più: erano descrizioni che *erano state date* della dipendenza tra situazioni e bandiere, e la dipendenza veniva mantenuta in parte per il fatto che gli agenti potevano consultare quella descrizione per sapere come comportarsi.

In generale, dev'esserci un qualche meccanismo intermedio che fa dipendere in modo definito le azioni di un agente dalle sue osservazioni. Il meccanismo deve essere sotto il controllo dell'agente in misura sufficiente a rendergli possibile regolarlo, così da produrre la dipendenza che desidera. Un meccanismo del genere è una descrizione della dipendenza desiderata che deve essere consultata dall'agente. Ma nell'esempio in questione non abbiamo motivo di pensare che il meccanismo generatore di dipendenza dell'aiutante e del guidatore contenga una descrizione della dipendenza desiderata o qualcosa di simile.

I nostri primi due esempi illustravano la specie standard di segnalazione, in versione formale e informale. Il terzo comporta una leggera modificazione; il quarto, il quinto e il sesto modificazioni più sensibili.

(3) Un uomo indica la strada facendo dei segni che dipendono da quella che egli ritiene sia la via migliore da seguire da parte degli altri. Le sue azioni vengono compiute una volta per tutte, ma le loro tracce rimangono a lungo. Quando gli altri, in seguito, seguono la sua strada, la via che prendono dipende dalle tracce che osservano delle sue azioni, cosicché essi prendono la via che egli ha ritenuto migliore per loro.

Negli esempi precedenti avevamo sequenze di problemi di coordinazione indipendenti tra coppie comunicante-uditorio. In questo esempio, ogni

traccia produce molti problemi di coordinazione a due persone aventi in comune colui che indica la strada. Le numerose persone che seguono le tracce devono tutte scegliere i loro piani circostanziali in modo da raggiungere la coordinazione con la scelta iniziale di un piano circostanziale da parte di colui che indicava la strada. Colui che indica la strada deve scegliere il suo piano in modo da raggiungere la coordinazione con ciascuna - o con quante più è possibile - delle scelte di piani compiute in occasioni diverse da parte di coloro che seguono le tracce. Ci sono tanti problemi di coordinazione diversi quante sono le occasioni in cui viene seguita la traccia. Questi problemi di coordinazione si estendono nel tempo, cosicché coloro che seguono le tracce devono disporre di piani circostanziali per rispondere all'osservazione delle tracce delle azioni di colui che ha segnato la strada, non all'osservazione delle azioni stesse.

(4) Una ferrovia installa dei segnali automatici, ossia dei semafori dotati di un congegno che fa dipendere la loro posizione dal fatto che il binario sia o meno occupato nel tratto successivo. Invece di un comunicante che compie azioni osservabili in base a un piano circostanziale, c'è un agente originario che installa il congegno, e c'è il congegno che in seguito funziona in base a un piano circostanziale. Come nell'esempio (3), c'è un numero qualunque di problemi di coordinazione a due persone con una parte in comune, e cioè la parte dell'agente che ha scelto il piano circostanziale da inserire nel congegno. Ma in questo esempio l'agente che sceglie un piano per il congegno di segnalazione non agisce egli stesso in accordo con quel piano.

Oppure i treni che si fermano e proseguono in risposta ai semafori potrebbero venire automatizzati. Su una ferrovia con treni automatici e semafori manuali, ogni agente che fa, funzionare un semaforo è implicato in un problema di coordinazione a due persone in cui l'agente ha scelto, una volta per tutte, un piano circostanziale da inserire nei treni. (Su una ferrovia avente sia treni che semafori automatizzati, c'è soltanto un problema di coordinazione singolo tra l'agente che sceglie un piano circostanziale da inserire nei treni e l'agente che sceglie un piano circostanziale da inserire nei semafori.)

(5) Un semaforo automatico viene installato in un crocicchio per regolare il traffico. La luce è sempre rossa su una strada e verde sull'altra. Cambia periodicamente, non in base a un piano circostanziale. I guidatori che si trovano (in qualunque momento) al crocicchio, invece, *agiscono* in base a un piano circostanziale convenzionale: fermarsi con il rosso, andare con il verde. Perciò essi oltrepassano l'incrocio con facilità e senza pericolo. I guidatori

- l'uditorio del semaforo - raggiungono un equilibrio di coordinazione. (Ancora una volta, gli incentivi indipendenti forniti dalla pattuglia della polizia stradale sono superflui se concordano con la convenzione, sono vani se non concordano, e quindi non rendono affatto meno convenzionale la regolarità.)

In questo caso la coordinazione riguarda esclusivamente i membri dell'uditorio. Nessuno svolge il ruolo di comunicante, né il semaforo né chi l'ha installato. Nessuno ha scelto un piano circostanziale per il semaforo; esso non ne ha. Questa è appunto una caratteristica della situazione che i guidatori possono sfruttare per mediare la loro coordinazione. È inessenziale che qualcuno sia responsabile della sua presenza: se costruiamo semafori perenni e i nostri discendenti dimenticassero che sono prodotti artificiali, le luci potrebbero regolare ancora il traffico e sarebbero ancora segnali. Anche il sorgere della luna può essere un segnale - di inizio di un'insurrezione, per esempio - anche se si tratterebbe di un segnale convenuto, valido in una sola occasione, non di un segnale convenzionale.

(6) Molte persone si alternano nell'uso degli stessi cavalli. Quando vogliono far girare a destra l'animale gridano "gii"; quando vogliono farlo girare a sinistra gridano "auu". Quando un cavallo non gira a destra al "gii" o a sinistra all'"auu", lo colpiscono. Gli uomini sono i comunicanti e i cavalli l'uditorio, ed entrambi agiscono in base a piani circostanziali complementari. Ma il problema di coordinazione consistente nello scegliere i piani e la convenzione che lo risolve esistono soltanto per gli uomini. I cavalli sono solo bestie, e reagiscono come sono stati addestrati a reagire. Gli uomini devono coordinarsi al fine di mantenere addestrati i cavalli e per trarre vantaggio dal loro addestramento.

## 4.2 Analisi della segnalazione

Adesso che abbiamo fatto conoscenza con i problemi di segnalazione, con i sistemi di segnalazione e con le convenzioni di segnalazione, come li definiremo?

In questo paragrafo limiterò la mia discussione alla segnalazione *a due parti*, in cui la coordinazione richiesta è una coordinazione tra comunicante e uditorio. Erano a due parti l'esempio di Paul Revere e del sagrestano e i primi tre esempi di convenzioni di segnalazione. Gli ultimi tre esempi illustravano la segnalazione *a una parte*, in cui la coordinazione richiesta era o tra i comunicanti o tra i membri dell'uditorio. Dopo aver definito i problemi, i sistemi e le convenzioni a due parti, sarebbe immediato ma noioso

dare definizioni analoghe di entrambi i tipi di segnalazione a una parte e della segnalazione in cui è richiesta una mescolanza di coordinazione a due parti e di coordinazione a una parte. Affiderò agli esempi l'illustrazione di come sarebbero tali definizioni. Ma quello che si dirà d'ora in avanti della segnalazione si intende applicato, *mutatis mutandis*, alla segnalazione in generale.

Un *problema di segnalazione a due parti* è una situazione  $S$  che comprende un agente detto *comunicante* e uno o più agenti detti *uditore*, e tale che è vero ed è conoscenza comune per il comunicante e per l'uditore che:

Vale esattamente uno fra i vari stati di cose alternativi  $s_1 \dots s_m$ .

Il comunicante, ma non l'uditore, è nella condizione migliore per dire di quale stato di cose si tratta.

Ciascuno fra i membri dell'uditore può compiere una qualunque tra varie azioni alternative  $r_1 \dots r_m$  dette *risposte*. Ciascuno vuole che le risposte dell'uditore dipendano in un certo modo dallo stato di cose che vale. C'è una certa funzione biunivoca  $F$ , da  $\{s_i\}$  su  $\{r_i\}$ , tale che ciascuno preferisce che ogni membro dell'uditore, per ogni  $s_i$ , faccia  $F(s_i)$  a condizione che valga  $s_i$ .

Il comunicante può compiere una qualunque tra varie azioni alternative  $\sigma_1 \dots \sigma_n$  ( $n \geq m$ ) dette *segnali*. L'uditore è nella condizione migliore per dire quale azione compie. Nessuno ha preferenze relative a queste azioni forti abbastanza da superare la sua preferenza per la dipendenza  $F$  delle risposte dell'uditore dagli stati di cose.

Si noti che viene specificato che la risposta preferita è la stessa per tutti i membri dell'uditore. Ciò può sembrare restrittivo, ma non è così. Se l'azione preferita di un membro dell'uditore dipende dalle condizioni in cui si trova, dal suo ruolo nella situazione, o da qualunque altra cosa oltre che dallo stato di cose in  $\{s_i\}$ , tale dipendenza va inserita nella specificazione delle risposte  $\{r_j\}$ . Se una nave da guerra è in pericolo, la risposta preferita di un'altra nave da guerra può essere questa: sostituirla nel combattimento se è effettivamente possibile farlo; altrimenti accorrere in suo aiuto. Noi consideriamo questa come una risposta, specificata da una coppia di condizionali.

Il *piano circostanziale di un comunicante* è un qualunque possibile modo in cui il segnale del comunicante può dipendere dallo stato di cose che egli osserva che vale. È una funzione  $F_c$  da  $\{s_i\}$  in  $\{\sigma_j\}$ . Un comunicante in  $S$  agisce in base a  $F_c$  se, per ogni  $s_i$ , compie  $F_c(s_i)$  se osserva che  $s_i$  vale. Poiché è in condizione di dire quale  $s_i$  vale, e poiché  $F_c(s_i)$  è la sua azione, egli

dovrebbe essere in grado di agire in base a qualunque piano circostanziale  $F_c$ . Se  $F_c$  è una funzione biunivoca, la chiamiamo *ammissibile*.

Analogamente, il *piano circostanziale di un uditorio* è un qualunque possibile modo in cui la risposta di un membro dell'uditorio può dipendere dal segnale che egli osserva provenire dal comunicante. È una funzione biunivoca  $F_a$  da una parte di  $\{\sigma_k\}$  in  $\{r_j\}$ . Un membro dell'uditorio in  $S$  agisce in base a  $F_a$  se, per ogni  $\sigma_k$  compreso nel dominio di  $F_a$ , compie  $F_a(\sigma_k)$  se osserva che il comunicante emette  $\sigma_k$ . Poiché è in condizione di dire quale  $\sigma_k$  è emesso, e poiché  $F_a(\sigma_k)$  è la sua azione, egli dovrebbe essere in grado di agire in base a qualunque piano circostanziale  $F_a$ . Se il codominio di  $F_a$  coincide con quello di  $F$ , chiamiamo  $F_a$  *ammissibile*.

Supponiamo che il comunicante agisca in base a un piano circostanziale  $F_c$  e che tutti i membri dell'uditorio agiscano in base a un piano circostanziale  $F_a$ . La risposta dell'uditorio dipenderà allora dallo stato di cose attuale secondo la funzione  $F_a|F_c$ , ottenuta per composizione di  $F_a$  e  $F_c$ . (Per due funzioni qualunque  $f$  e  $g$ , e per qualunque argomento  $x$  tale che  $g(x)$  è nel dominio di  $f$ ,  $f|g(x)$  è definita come  $f(g(x))$ ; per altri argomenti  $f|g$  è indefinita.)

Se lo stato di cose attuale è  $s$ , e se  $F_a|F_c(s) = F(s)$ , la risposta dell'uditorio sarà una risposta preferita nello stato di cose attuale. Così ciascun agente agirà in base a uno dei migliori piani circostanziali, dati i piani circostanziali degli altri e lo stato di cose attuale.

Meglio, supponiamo che  $F_a|F_c = F$  uniformemente su  $\{s_i\}$ . (Ne segue che il codominio di  $F_c$  deve coincidere con il dominio di  $F_a$ .) La risposta dell'uditorio sarà allora una risposta che è preferita nello stato di cose attuale, indipendentemente da quello che può essere lo stato di cose in  $\{s_i\}$ . Ciascun agente agirà in base a uno dei migliori piani circostanziali, dati i piani circostanziali degli altri e *qualunque* stato di cose. Ogniqualvolta  $F_c$  e  $F_a$  si combinano in questo modo per produrre la dipendenza della risposta dell'uditorio dallo stato di cose che viene preferita, chiamiamo  $\langle F_c, F_a \rangle$  *sistema di segnalazione*.

Tutti e solo i piani circostanziali ammissibili appartengono ai sistemi di segnalazione. Dimostrazione:

Sia  $F_c$  il piano circostanziale di un comunicante. Se è ammissibile, è una funzione biunivoca e ha un'inversa  $F_c^{-1}$ ; in questo caso,  $\langle F_c, F_c^{-1} \rangle$  è un sistema di segnalazione. Se è inammissibile non è una funzione biunivoca; così, per qualunque  $F_a$ , o  $F_a|F_c$  non è una funzione biunivoca oppure  $F_a|F_c$  non è definita per

qualche stato di cose; in nessun caso  $F_a|F_c$  può coincidere con  $F$ ; così  $\langle F_c, F_a \rangle$  non è un sistema di segnalazione.

Sia  $F_a$  il piano circostanziale di un uditorio. Se è ammissibile, sia  $F\check{a}$  la sua inversa;  $\langle F\check{c}|F, F_a \rangle$  è un sistema di segnalazione. Se è inammissibile il suo codominio non coincide con quello di  $F$ ; così, per qualunque  $F_c$ , il codominio di  $F_a|F_c$  non coincide con quello di  $F$ ; così  $\langle F_c, F_a \rangle$  non è un sistema di segnalazione.

In un problema di segnalazione con  $m$  stati di cose e  $n$  segnali, ci sono  $n!/(n - m)!$  sistemi di segnalazione. Dimostrazione:

Costruiamo  $F_c$  nel modo seguente: per qualunque  $s$  in  $\{s_i\}$ , se  $F(s)$  è il  $k$ -esimo membro del codominio di  $F$ , allora  $F_c(s)$  sia  $\sigma_k$ .  $F_c$  è un piano circostanziale ammissibile del comunicante. Prendiamo una qualunque funzione  $G$  che applica biunivocamente il codominio di  $F_c$  in  $\sigma_k$ . Sia  $\check{G}$  l'inversa di  $G$ ; sia  $F\check{c}$  l'inversa di  $F_c$ . Allora  $\langle G|F_c, (F|F\check{c})|\check{G} \rangle$  è un sistema di segnalazione. Ogni sistema di segnalazione può essere ottenuto in questo modo. Ci sono  $n!/(n - m)!$  diverse funzioni biunivoche dal codominio' di  $F_c$  in  $\sigma_k$ , per cui ci sono altrettanti sistemi di segnalazione diversi.

Possiamo pensare che il comunicante e l'uditorio agiscano prima scegliendo dei piani circostanziali e poi agendo in base ai piani da loro scelti. Se è così, le loro scelte di piani costituiscono un problema di coordinazione. Possiamo rappresentarlo per mezzo di una matrice in cui le righe rappresentano i piani circostanziali del comunicante e le colonne, i livelli ecc., rappresentano i piani circostanziali di tutti i membri dell'uditorio. Alcuni di questi piani sono ammissibili, altri no. La vincita di un agente relativa a una combinazione dipende dalle sue vincite relative a quella combinazione in ciascuno degli stati di cose in  $\{s_i\}$ , più le probabilità che egli assegna ai diversi stati di cose. (Per alcune combinazioni, dovremo considerare quello che egli si aspetta che i membri dell'uditorio facciano in risposta a un segnale che non appartiene ai domini dei piani da loro scelti.) È facile vedere come, in questo problema di coordinazione, ogni sistema di segnalazione sia un equilibrio di coordinazione proprio.

Ma i sistemi di segnalazione possono non essere i soli equilibri di coordinazione. Certe coppie di piani circostanziali inammissibili possono essere equilibri di coordinazione. Supponiamo, nel nostro esempio di Paul Revere e del sagrestano, che la difesa contro un attacco via terra funzioni abbastanza bene contro un attacco via mare, ma non viceversa, e supponiamo che

i due attacchi sembrano equiprobabili. Allora ci può essere un equilibrio di coordinazione tra questo piano del sagrestano:

Se si scorge che i soldati inglesi non si spostano, non appendere lanterne nella cella campanaria.

Se si scorge che i soldati inglesi si spostano via terra, appendere una lanterna nella cella campanaria.

Se si scorge che i soldati inglesi si spostano via mare, appendere una lanterna nella cella campanaria.

e questo piano di Paul Revere:

Se non si scorge alcuna lanterna appesa nella cella campanaria, tornare.

Se si scorge una lanterna appesa nella cella campanaria, avvisare la popolazione che i soldati inglesi arrivano via terra.

I piani si combinano in modo da produrre una dipendenza della risposta di Paul Revere dallo stato di cose osservato, la quale, pur non essendo la dipendenza preferita, non è tuttavia troppo cattiva. Ma nessuno dei due può raggiungere la dipendenza preferita semplicemente cambiando il proprio piano; ciascuno agisce in base a un piano che, dato quello dell'altro, è il migliore. Così la combinazione è un equilibrio di coordinazione. (Un equilibrio improprio; dato il piano del sagrestano, Paul Revere avrebbe potuto altrettanto bene scegliere un piano diverso:

Se non si scorge alcuna lanterna appesa nella cella campanaria, tornare.

Se si scorge una lanterna appesa nella cella campanaria, avvisare la popolazione che i soldati inglesi arrivano via terra.

Se si scorgono due lanterne appese nella cella campanaria, avvisare la popolazione che i soldati inglesi arrivano via mare.

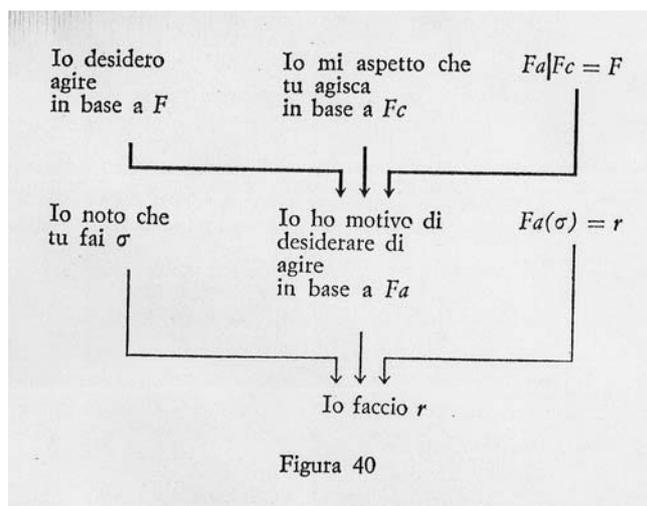
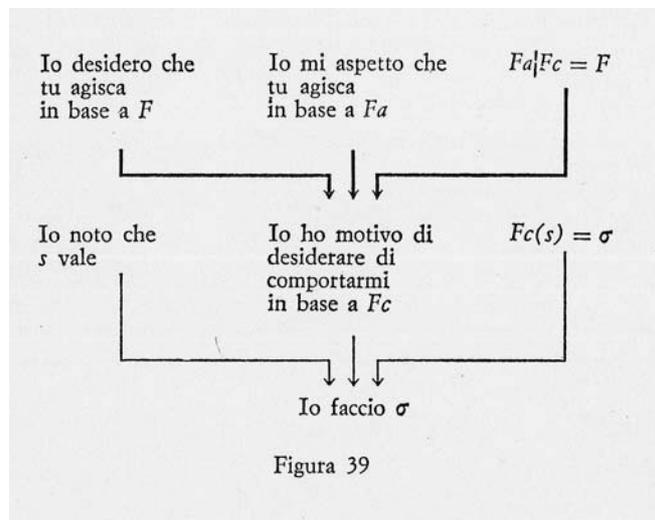
La sua risposta sarebbe stata esattamente la stessa, indipendentemente da quello che il sagrestano avesse scorto.) Non si tratta di un sistema di segnalazione. Perlomeno, non è un sistema di segnalazione nel nostro problema di segnalazione *iniziale*. Se classifichiamo di nuovo gli stati di cose e le risposte, trascurando la differenza tra le due specie di attacco e le due specie di avviso, otteniamo un nuovo problema di segnalazione con una funzione di dipendenza preferita meno dettagliata; e, in questo nuovo problema, la combinazione di piani che stiamo considerando è un sistema di segnalazione e un equilibrio di coordinazione proprio.

Nei problemi di segnalazione abbiamo quindi due specie di equilibri di coordinazione. Ci sono i sistemi di segnalazione; e ci sono equilibri di coordinazione impropri che non sono sistemi di segnalazione ma che lo diventerebbero se la classificazione degli stati di cose e delle risposte fosse meno dettagliata. Non so se ci possano essere equilibri di coordinazione di altre specie. Un risultato pertinente è questo: se l'uditorio ha esattamente un membro, nessuna combinazione di un piano ammissibile con un piano inammissibile può essere un equilibrio; infatti, un'altra combinazione sulla stessa riga o sulla stessa colonna sarebbe un sistema di segnalazione e sarebbe preferita alla combinazione data.

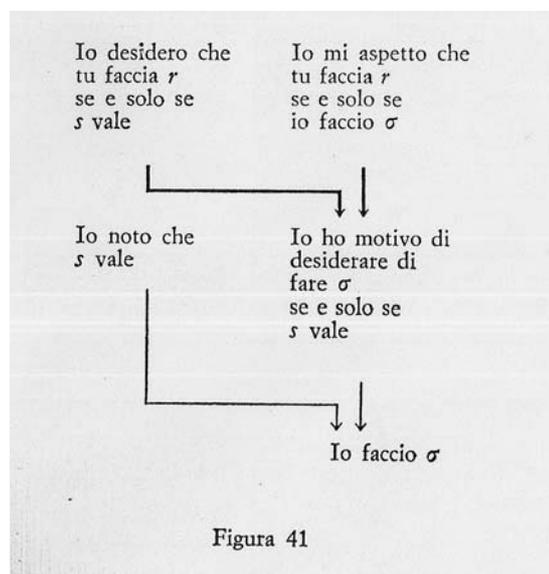
Data la definizione di sistema di segnalazione, possiamo definire una *convenzione di segnalazione* come una qualunque convenzione per mezzo della quale i membri di una popolazione P implicati, come comunicanti o come uditorio, in un certo problema di segnalazione svolgono la loro parte in un certo sistema di segnalazione  $\langle Fc, Fa \rangle$  agendo in base ai loro rispettivi piani circostanziali. Se una tale convenzione esiste, chiameremo anche  $\langle Fc, Fa \rangle$  *sistema di segnalazione convenzionale*.

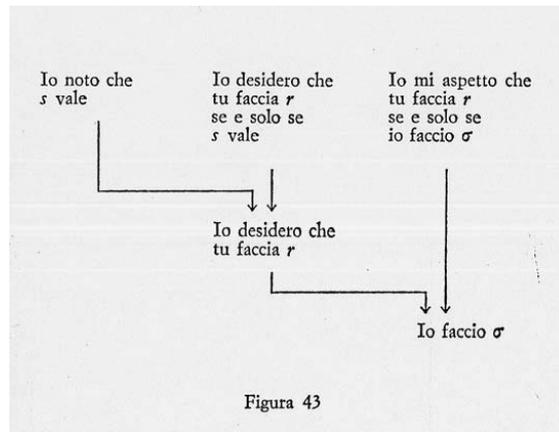
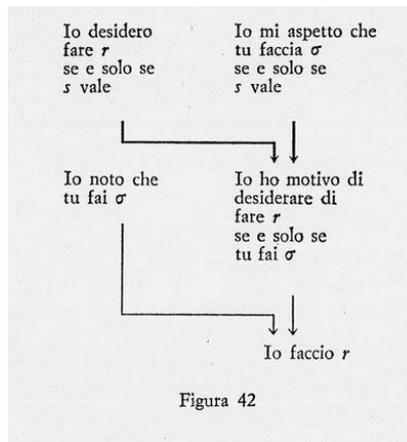
Se risulta che alcuni problemi di segnalazione hanno degli equilibri di coordinazione che non sono sistemi di segnalazione - neppure nell'ipotesi di un'opportuna classificazione meno dettagliata degli stati di cose e delle risposte - allora dovremmo dire che ci possono essere delle convenzioni che non sono convenzioni di segnalazione anche se regolano la scelta di piani circostanziali in un problema di segnalazione. Ma non vedo nessun male nel lasciare aperta questa possibilità.

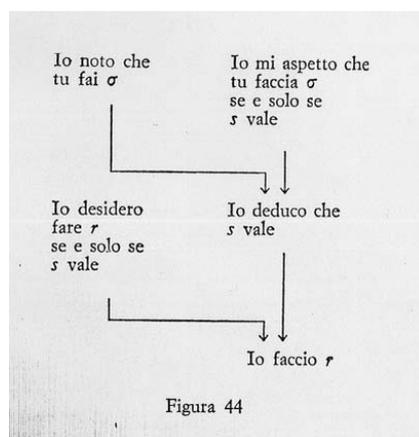
Quando un agente si comporta in base a una convenzione di segnalazione - e in qualunque altro caso in cui egli risolve un problema di segnalazione per mezzo dell'accordo, della salienza o del precedente - la sua azione è giustificata dal ragionamento illustrato nella Figura 39 (per il comunicante) o nella Figura 40 (per un membro dell'uditorio). In altri termini, stiamo trattando la scelta di un'azione da parte di un agente come scelta razionale di un intero piano circostanziale, seguita dall'azione intesa a realizzare il suo piano alla luce di informazioni particolari relative alla situazione. È questa la trattazione migliore se vogliamo far entrare la segnalazione nella teoria generale della coordinazione e della convenzione.



Ma potremmo trattare in modo più economico le giustificazioni che gli agenti hanno per le proprie scelte. L'informazione particolare di cui l'agente ha bisogno - quale stato di cose vale o quale segnale è stato emesso - gli è sempre disponibile; così la sua scelta è giustificata se egli giustifica soltanto la parte pertinente del suo piano circostanziale. Scartando le circostanze non pertinenti, otteniamo il ragionamento illustrato nella Figura 41 (per un comunicante) o nella Figura 42 (per membro dell'uditorio). Queste giustificazioni possono essere leggermente riordinate, in modo da eliminare anche la scelta di una parte del piano circostanziale. Le premesse disponibili possono essere applicate in un ordine diverso. Otteniamo così, per il comunicante, il ragionamento illustrato nella Figura 43. Per un membro dell'uditorio, otteniamo il ragionamento illustrato nella Figura 44.







Nel caso del comunicante, lo stato di cose  $s$  gli dà motivo di cercare di produrre la risposta dell'uditorio  $r$ ; e la sua aspettativa che l'uditorio risponderà a  $\sigma$  in base a  $Fa$  gli dà motivo di credere che  $\sigma$  sia un mezzo efficace per produrre  $r$ . Così egli fa  $\sigma$ . Nel caso di un membro dell'uditorio, la sua aspettativa che il comunicante agisca in base a  $Fc$  gli dà motivo di considerare  $\sigma$  come una buona prova del fatto che  $s$  vale; e, avendo dedotto che  $s$  vale, egli ha motivo di fare  $r$ .

Questi nuovi schemi di giustificazione, come quelli precedenti, possono venire ricostruiti; la ricostruzione, da parte di un agente, del ragionamento dell'altro gli fornisce la premessa centrale del proprio ragionamento. Ma queste ricostruzioni devono essere condizionali: un agente calcola in che modo *ragionerebbe* l'altro se gli venisse proposta una certa premessa. Il comunicante calcola come ragionerebbe l'uditorio in risposta a un segnale; l'uditorio calcola come ragionerebbe il comunicante in risposta a uno stato di cose. Queste ricostruzioni condizionali possono essere a loro volta ricostruite e così via, per un qualunque numero di passi. Consideriamo il ragionamento del comunicante che giustifica la sua decisione di fare  $\sigma$ ; possiamo farlo regredire di due passaggi, come è illustrato nella Figura 45.

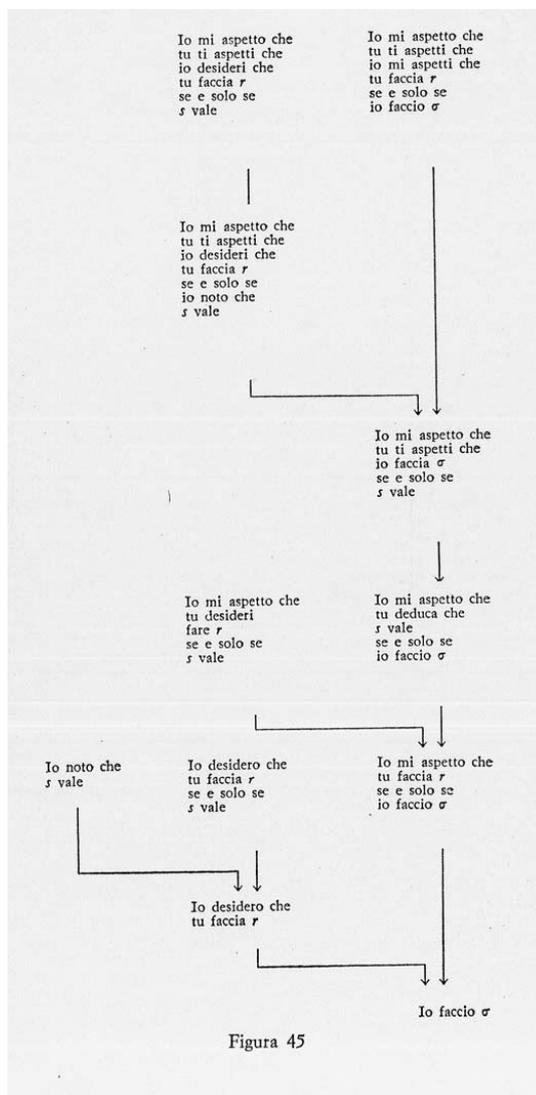


Figura 45

Abbiamo rappresentato le giustificazioni degli agenti in due modi completamente diversi. Il primo è basato sul fatto che le scelte di un piano circostanziale da parte dell'agente costituiscono un problema di coordinazione, relativamente al quale il sistema di segnalazione convenzionale  $\langle Fc, Fa \rangle$  è un equilibrio di coordinazione. Il secondo modo è basato sul fatto che, siccome è conoscenza comune che l'uditorio agisce in base a  $Fa$ , è conoscenza comune che un segnale  $\sigma$  è un buon mezzo per produrre la risposta  $Fa(\sigma)$ ; e sul fatto che, siccome è conoscenza comune che il comunicante agisce in

base a  $F_c$ , è conoscenza comune che un segnale  $\sigma$  è una buona prova del fatto che vale lo stato che è applicato su  $\sigma$  da  $F_e$ .

Tuttavia, entrambe le rappresentazioni sono in qualche modo artificiali, in quanto rappresentazioni del ragionamento *effettivo* di un agente. Fortunatamente, sono artificiali in modi opposti e complementari. Se l'una è realistica, l'altra non lo è. Prendiamo un agente che non si limita a seguire ciecamente il proprio piano circostanziale. Dopo che ha saputo qual è la situazione attuale, egli riflette sugli esiti probabili delle sue azioni alternative e compie quella che gli sembra migliore. Dovremo veramente dire che egli sceglie - o sceglie di nuovo - l'intero piano circostanziale? O soltanto la parte che riguarda la situazione attuale? Prendiamo, d'altra parte, un agente la cui segnalazione è un noioso lavoro di routine, svolto con la mente altrove. Forse che egli si sofferma a pensare (se è un comunicante): "Questa azione indurrà l'uditorio a comportarsi così-e-così, il che è il meglio nella situazione attuale; quindi sarà bene che io compia questa azione", oppure (se fa parte dell'uditorio): "Il comunicante deve comportarsi così perché ha osservato questo-e-questo; in tal caso l'azione che sarebbe meglio che io compissi sarebbe questa, per cui farò bene a compiere questa azione"? No: egli si limita a scegliere il suo piano circostanziale per abitudine. È consapevole, in generale, del fatto che il piano circostanziale che sta seguendo è il migliore, dato il piano che sta seguendo il suo partner. Se ricevesse delle prove in senso contrario, si fermerebbe a pensare e comincerebbe a agire in base a un nuovo piano circostanziale. Ma finché tutto va bene, egli non pensa a quale possa essere il suo comportamento migliore in qualunque occasione particolare.

Ma non abbiamo bisogno di rappresentarci l'effettivo ragionamento degli agenti. È sufficiente che consideriamo le giustificazioni razionali delle loro scelte per mezzo del ragionamento pratico che essi potrebbero sviluppare, a partire dalle loro credenze e dai loro desideri. Ma ciò non significa rinunciare all'interesse di spiegare le loro scelte. Le giustificazioni spiegano effettivamente le scelte, che l'agente sviluppi effettivamente o meno un processo ragionato per cogliere la giustificazione. Infatti, è connaturato alla natura umana che noi tendiamo ad agire in un modo giustificato dalle nostre credenze e dai nostri desideri, anche quando non cogliamo la giustificazione. Per esprimermi in forma negativa: quali che possano essere i processi abituali che controllano effettivamente le nostre scelte, se essi cominciassero a contrastare con le nostre credenze e i nostri desideri verrebbero ben presto interrotti, corretti e ricondizionati dal nostro ragionamento pratico esplicito.

### 4.3 Segnalazione verbale

Un'azione è adatta a essere un segnale se essa o le sue tracce sono facili da osservare da parte di un uditorio; e se è intrinsecamente irrilevante, in modo che nessuna delle parti abbia forti preferenze estranee pro o contro la sua esecuzione.

Due specie di azioni sono adatte in modo ideale a fungere da segnali, in molte situazioni: la produzione di una successione di suoni vocali entro l'area in cui si può venir uditi dall'uditorio; e la produzione di una successione di caratteri durevoli su una superficie liscia che sia visibile all'uditorio. In breve, il parlare e lo scrivere.

Definiamo *espressione verbale* qualunque successione finita di tipi di suoni vocali o di tipi di caratteri.<sup>1</sup> Emettere o scrivere un'espressione verbale è produrre una successione di suoni che si propagano nell'aria o una successione di caratteri su una superficie. Un segnale che consiste di un'azione di emissione o di iscrizione di un'espressione verbale può essere chiamato *segnale verbale*.

Ufficialmente, il segnale è un'azione. Ma di solito non è necessario che facciamo attenzione a distinguere l'azione, la successione di suoni o di caratteri prodotta dall'azione, e l'espressione verbale emessa o scritta; tutte e tre possono venir chiamate segnale. Possiamo anche permettere che "emissione" comprenda sia la scrittura sia l'emissione vocale.

Se le circostanze fossero state diverse, il sagrestano e Paul Revere avrebbero potuto trovare la soluzione migliore in un sistema di segnalazione verbale convenuto o convenzionale. Per segnalare che scorgeva i soldati inglesi spostarsi via mare, il sagrestano avrebbe potuto sussurrare dei suoni all'orecchio di Paul Revere o mandargli un pezzo di carta con sopra scritti dei caratteri. Questi suoni o caratteri potrebbero essere uno speciale codice convenuto, oppure potrebbero essere "I soldati inglesi arrivano via mare". In quest'ultimo caso, la segnalazione verbale del sagrestano sarebbe indistinguibile dall'uso ordinario del linguaggio.

Se provvediamo una ipotetica comunità di un gran numero di convenzioni di segnalazione verbale da usare in varie attività, e di espressioni verbali opportunamente scelte *ad hoc*, saremo in grado di simulare abbastanza bene una comunità di utenti linguistici: cioè noi stessi. Un osservatore che stesse sullo sfondo a osservare queste persone che, mentre sbrigano i loro affari, usano i segnali verbali convenzionali, potrebbe metterci parecchio tempo a ren-

---

<sup>1</sup>In base a questa definizione può esistere un'espressione verbale che non è mai emessa. Cfr. Quine, *Word and Object*, cit., pp. 194-195. [Cfr. trad. it. cit., p. 239. (N.d.T.)]

dersi conto che non si tratta di normali utenti linguistici. Ma un osservatore che cercasse di conversare con loro noterebbe alcune deficienze. Scoprirebbe che ogni espressione verbale che essi impiegano è associata convenzionalmente a qualche stato di cose facilmente osservabile, o a qualche definita azione di risposta, o ad ambedue. E scoprirebbe che essi possono usare espressioni verbali soltanto in numero finito, in modo che le convenzioni che regolano la loro segnalazione verbale possano essere descritte menzionando ciascuna delle espressioni che impiegano.

Vedremo in seguito come la segnalazione verbale non esaurisca l'intero uso del linguaggio. Tuttavia rimane vero che i nostri ipotetici emittenti di segnali verbali non fanno niente che noi non facciamo. Soltanto, noi facciamo di più. Il loro uso del linguaggio duplica un frammento del nostro linguaggio.

Quell'uso è interamente regolato dalle convenzioni di segnalazione. Che dire del nostro, in quelle occasioni in cui è esattamente uguale al loro? Non vedo alcuna ragione per negare che anche noi abbiamo delle convenzioni di segnalazione verbale. Ciò che *sarebbe* segnalazione convenzionale se fosse il nostro unico uso del linguaggio non ha caratteristiche diverse per il fatto che noi usiamo il linguaggio anche in altri modi. Esaminando la definizione di convenzione di segnalazione verbale, possiamo verificare che è soddisfatta da una parte del nostro uso del linguaggio.

Ma anche se alcune delle convenzioni che regolano questo uso sono convenzioni di segnalazione verbale, ciò non implica che tali convenzioni comparirebbero in una specificazione completa delle nostre convenzioni di linguaggio. Probabilmente verrebbero incluse in convenzioni più generali. Ricordiamo, dal capitolo 2.5, che le particolarizzazioni di convenzioni possono essere convenzioni a pieno diritto. Una convenzione generale di linguaggio, che regola espressioni verbali diverse in numero infinito e vale in un'ampia gamma di situazioni, avrebbe conseguenze limitate a una classe finita di espressioni e a situazioni strettamente specifiche. Le nostre convenzioni di segnalazione verbale potrebbero essere conseguenze di questo tipo.

#### 4.4 Significato convenzionale dei segnali

Ho cercato di dimostrare che una spiegazione adeguata della segnalazione non deve necessariamente menzionare il significato dei segnali: almeno, non per nome. Ma naturalmente i segnali hanno significato.

Supponiamo che  $\langle Fc, Fa \rangle$  sia un sistema di segnalazione convenzionale;  $\sigma$  è un segnale che gli appartiene;  $s$  è lo stato di cose applicato su  $\sigma$  da  $Fc$ ;  $r$  è la risposta su cui  $\sigma$  è applicato da  $Fa$ .

Potremmo allora chiamare  $\sigma$  - in generale e in qualunque occasione particolare in cui venga emesso in conformità alla convenzione - un *segnale convenzionale che s vale*; e potremmo dire che  $\sigma$  *significa convenzionalmente che s vale*. Infatti  $\sigma$  sta a provare che s vale, in virtù del fatto che il comunicante agisce in base a Fc, e quindi in virtù della convenzione. Oppure potremmo chiamare  $\sigma$  un *segnale convenzionale di fare r*; e potremmo dire che *significa convenzionalmente di fare r*. Infatti è un ottimo modo di ottenere che l'uditorio faccia r, in virtù del fatto che esso agisce in base a Fa, e quindi in virtù della convenzione.

In che modo dovremmo rappresentare il significato di  $\sigma$  : come un segnale-che o come un segnale-di? A volte possiamo rappresentarlo adeguatamente in entrambi i modi: per esempio, in tutti gli esempi che abbiamo fatto finora. Ma in alcuni casi è appropriato soltanto descrivere  $\sigma$  come un segnale che uno stato di cose vale; in questi casi chiamiamo  $\sigma$  *segnale indicativo* (nel sistema di segnalazione in questione). In altri casi è appropriato soltanto descrivere  $\sigma$  come segnale di fare qualcosa; in questi casi chiamiamo  $\sigma$  *segnale imperativo*. (Chiamiamo  $\sigma$  segnale neutro se lo si può chiamare altrettanto propriamente segnale-che o segnale-di.) Sembra che la differenza stia nel carattere dei piani Fa e Fe. Quale grado di libertà lasciano questi al comunicante o all'uditorio?

Un piano circostanziale può essere o non essere discrezionale, cioè può richiedere o meno che un agente decida quale linea d'azione è la migliore per lui e per i suoi partner. Suggestisco che, se Fa è discrezionale e Fc non lo è, allora  $\sigma$  è *indicativo*; se Fc è discrezionale e Fa non lo è,  $\sigma$  è *imperativo*; se nessuno dei due o entrambi sono discrezionali, allora  $\sigma$  è *neutro*.

Consideriamo il segnale del sagrestano consistente nell'appendere due lanterne nella cella campanaria. (Fingeremo che il sistema di segnalazione sia un sistema convenzionale, non limitato a una sola occasione.) Se il sistema di segnalazione è  $\langle Fe, Fa \rangle$  con

- Fc: Se si scorge che i soldati inglesi non si spostano, non appendere lanterne nella cella campanaria.  
 Se si scorge che i soldati inglesi si spostano via terra, appendere una lanterna nella cella campanaria.  
 Se si scorge che i soldati inglesi si spostano via mare, appendere due lanterne nella cella campanaria.
- Fa: Se non si scorge alcuna lanterna appesa nella cella campanaria, tornare.  
 Se si scorge una lanterna appesa nella cella campanaria, avvisare la popolazione che i soldati inglesi arrivano via terra.

Se si scorgono due lanterne appese nella cella campanaria, avvisare la popolazione che i soldati inglesi arrivano via mare.

allora l'appendere due lanterne è il segnale che i soldati inglesi sono stati visti spostarsi via mare, ed è il segnale di avvisare la popolazione che i soldati inglesi arrivano via mare. Le due descrizioni sembrano ugualmente appropriate; il segnale è neutro.

Ma se il sistema di segnalazione è  $\langle Fc, Fa' \rangle$ , dove  $Fa'$  è una variante discrezionale di  $Fa$ ,

$Fa'$ : Se non si scorge alcuna lanterna appesa nella cella campanaria, fare ciò che sembra meglio nell'ipotesi che i soldati inglesi non siano stati visti spostarsi.

Se si scorge una lanterna appesa nella cella campanaria, fare ciò che sembra meglio nell'ipotesi che i soldati inglesi siano stati visti spostarsi via terra.

Se si scorgono due lanterne appese nella cella campanaria, fare ciò che sembra meglio nell'ipotesi che i soldati inglesi siano stati visti spostarsi via mare.

allora l'appendere due lanterne è definito propriamente come il segnale che i soldati inglesi sono stati visti spostarsi via mare. Sarebbe strano definirlo come il segnale di fare ciò che sembra meglio nell'ipotesi che i soldati inglesi siano stati visti spostarsi via mare. In  $\langle Fc, Fa' \rangle$ , l'appendere due lanterne è un segnale indicativo.

D'altra parte, se il sistema di segnalazione è  $\langle Fc', Fa \rangle$ , dove  $Fc'$  è una variante discrezionale di  $Fc$ ,

$Fc'$ : Se sembra che sia meglio che Paul Revere torni, non appendere lanterne nella cella campanaria.

Se sembra meglio che Paul Revere avvisi la popolazione che i soldati inglesi arrivano via terra, appendere una lanterna nella cella campanaria.

Se sembra meglio che Paul Revere avvisi la popolazione che i soldati inglesi arrivano via mare, appendere due lanterne nella cella campanaria.

allora l'appendere due lanterne è definito propriamente come il segnale di avvisare la popolazione che i soldati inglesi arrivano via mare. Sarebbe strano definirlo come il segnale che sembra meglio che Paul Revere avvisi la popolazione che i soldati inglesi arrivano via mare. In  $\langle Fc', Fa \rangle$ , l'appendere

due lanterne è un segnale imperativo.

Fa e Fa' sono formulati in termini di classificazioni diverse e incrociate dell'azione di Paul Revere. Fc e Fc' sono formulati in termini di classificazioni diverse e incrociate di stati di cose. Così i tre sistemi di segnalazione che abbiamo descritto appartengono a problemi di segnalazione diversi. Ciascuno è un sistema di segnalazione entro il proprio problema di segnalazione, in quanto fa sì che l'azione di Paul Revere dipenda dallo stato di cose nel miglior modo possibile entro le classificazioni date.

Ho trascurato  $\langle Fc', Fa' \rangle$  perché non sono sicuro che nel suo problema di segnalazione produca la dipendenza preferita; è verosimile anche che il sagrestano e Paul Revere si ostacolino a vicenda. Se è un sistema di segnalazione, i suoi segnali sono neutri.

Un agente che agisce in base a un piano circostanziale discrezionale deve far uso del proprio giudizio per rispondere in modo elastico alla situazione nel suo complesso quale egli la vede, così da essere in grado di scegliere una strategia particolare per affrontare le circostanze impreviste. Paul Revere, agendo in base a Fa' e vedendo due lanterne, potrebbe decidere che è più urgente nascondere i capi che avvertire la popolazione. Il sagrestano, agendo in base a Fc' e osservando che i soldati inglesi si spostano via mare, potrebbe scoprire che essi si dirigono nel luogo sbagliato; per cui sarebbe meglio non ostacolarli e aspettare un attacco successivo più importante. Il prezzo dell'elasticità è, naturalmente, il rischio dell'informazione sbagliata o del giudizio sbagliato.

Un agente potrebbe far uso del proprio giudizio anche nel caso che il suo sistema di segnalazione convenzionale gli offra un piano circostanziale non discrezionale. Potrebbe cioè decidere che è meglio violare la convenzione di segnalazione in vista di circostanze impreviste; e potrebbe avere ragione. Ma, se lo facesse, il suo esercizio del giudizio sarebbe inaspettato; mentre, se agisse in base a un piano discrezionale in conformità a una convenzione di segnalazione, sarebbe conoscenza comune che egli fa uso del proprio giudizio.

Nel definire il significato dei segnali non è affatto necessario che ci limitiamo ai sistemi di segnalazione convenzionale. Consideriamo un qualunque sistema di segnalazione  $\langle Fc, Fa \rangle$  per un problema di segnalazione S, che  $\langle Fc, Fa \rangle$  venga adottato convenzionalmente o meno in una popolazione qualsiasi. Se Fa è discrezionale e Fc no, i segnali di  $\langle Fc, Fa \rangle$  sono indicativi. Consideriamo un segnale  $\sigma$  e lo stato di cose s applicato da Fc su  $\sigma$ . Possiamo chiamare  $\sigma$  un segnale in  $\langle Fc, Fa \rangle$  che s vale; e possiamo dire che  $\sigma$  significa in  $\langle Fc, Fa \rangle$  che s vale. Se Fc è discrezionale e Fa no, i segnali di  $\langle Fc, Fa \rangle$  sono imperativi. Consideriamo un segnale  $\sigma$  e la risposta r su cui  $\sigma$  è applicato da Fa. Possiamo chiamare  $\sigma$  un segnale in

$\langle Fc, Fa \rangle$  di fare  $r$ ; e possiamo dire che  $\sigma$  significa in  $\langle Fc, Fa \rangle$  di fare  $r$ . Infine, se né  $Fc$  né  $Fa$  sono discrezionali, oppure lo sono entrambi, i segnali di  $\langle Fc, Fa \rangle$  sono neutri; possono essere chiamati altrettanto correttamente segnali-che o segnali-di in  $\langle Fc, Fa \rangle$ ; e si può dire altrettanto correttamente che significano-che o che significano-di in  $\langle Fc, Fa \rangle$ . D'ora in poi tralascerò quasi completamente i segnali neutri.

Ci aspettiamo che, dando il significato dei segnali indicativi, vengano date anche le loro condizioni di verità. E così avviene. Sia  $\sigma$  il segnale indicativo che  $s$  vale, in un sistema di segnalazione  $\langle Fc, Fa \rangle$  per il problema di segnalazione  $S$ . Allora possiamo chiamare  $\sigma$  vero in qualunque esempio di  $S$  in cui  $s$  vale, e falso in qualunque esempio di  $S$  in cui  $s$  non vale.

Ufficialmente, ricordiamo, il segnale  $\sigma$  è un'azione. Stiamo quindi attribuendo la verità e la falsità a delle azioni. Non abbiamo stabilito se i segnali sono tipi di azione o azioni particolari; né siamo tenuti a stabilirlo adesso. Il tipo-di-azione  $\sigma$  è vero o falso, in relazione a qualunque esempio di  $S$ , a seconda che lo stato di cose  $s$  valga o non valga in quell'esempio. Non abbiamo alcun motivo di limitarci a quegli esempi di  $S$  in cui  $\sigma$  viene eseguito. Un'azione particolare  $\sigma$  di quel tipo è vera o falsa a seconda che lo stato di cose  $s$  valga o non valga nell'esempio di  $S$  in cui  $\sigma$  viene eseguita. Se  $\sigma$  è un segnale verbale, possiamo indulgere a un'altra ambiguità innocua: possiamo dire che la successione di suoni o caratteri prodotta dall'azione  $\sigma$  in un esempio particolare di  $S$  è vera o falsa, e possiamo dire che l'espressione verbale emessa nell'azione  $\sigma$  è vera o falsa relativamente a un qualunque esempio di  $S$ . I filosofi hanno discusso a lungo sul problema di sapere quali entità siano propriamente portatrici della verità e delle altre proprietà semantiche; io adotterò un punto di vista eclettico, permettendo che portatrici di proprietà semantiche siano diverse specie di entità. Non vedo in ciò alcun pericolo, purché siamo pronti a convertire, ogniqualvolta è opportuno farlo, le proprietà semantiche in relazioni semantiche (per esempio tra espressioni verbali e esempi di  $S$ ).

Per definizione di problema di segnalazione e di sistema di segnalazione, in qualunque esempio di  $S$  uno dei segnali di  $\langle Fc, Fa \rangle$  è vero e gli altri sono falsi. Supponiamo che  $\langle Fc, Fa \rangle$  sia un sistema di segnalazione convenzionale in una popolazione. Possiamo allora descrivere nel modo seguente la regolarità convenzionale del loro comportamento: in qualsiasi esempio di  $S$ , il comunicante cerca di emettere qualunque segnale sia vero in quell'esempio, e ogni membro dell'uditorio risponde facendo tutto ciò che gli sembra meglio nell'ipotesi che il segnale del comunicante sia vero.

Ma non sembra che questa impostazione del problema sia giusta. Come può la veracità essere una convenzione? Qual è l'alternativa ad essa? La

non-veracità sistematica? Supponiamo di aver ordinato i segnali di  $\langle Fe, Fa \rangle$  in una qualche specie di ordine ciclico. I membri di P potrebbero allora cavarsela molto bene se, ogniqualvolta un segnale  $\sigma$  fosse vero, essi emettessero non  $\sigma$  bensì il segnale immediatamente successivo a  $\sigma$  nell'ordine (per esempio successivo a destra): un segnale falso. Questa specie di non veracità sistematica, se uniforme e attesa uniformemente, funzionerebbe altrettanto bene della veracità sistematica. Ma ciò avverrebbe perché si tratterebbe in realtà di veracità sistematica: non nel sistema di segnalazione iniziale  $\langle Fc, Fa \rangle$ , ma in un sistema di segnalazione alternativo ottenuto da  $\langle Fc, Fa \rangle$  permutando i segnali.

Ecco la soluzione: abbiamo sempre parlato non della verità in generale, ma della verità nel sistema di segnalazione  $\langle Fc, Fa \rangle$ . Avremmo dovuto formulare la convenzione così: in qualunque esempio di S fra membri di P, il comunicante cerca di emettere qualunque segnale sia vero in  $\langle Fc, Fa \rangle$  in quell'esempio, e l'uditorio risponde facendo tutto ciò che sembra meglio nell'ipotesi che egli sia riuscito nel suo intento. Quella che chiamavamo convenzione di veracità si chiama, con più accuratezza, convenzione di veracità in  $\langle Fc, Fa \rangle$ .

Data questa più accurata formulazione della convenzione, non abbiamo più difficoltà a scoprire le sue alternative. L'alternativa a una convenzione di veracità in  $\langle Fc, Fa \rangle$  sarebbe una regolarità di veracità in qualche altro sistema di segnalazione  $\langle Fc', Fa' \rangle$ . Quest'ultimo può essere un qualunque sistema di segnalazione per il problema S che sia abbastanza diverso da  $\langle Fc, Fa \rangle$ : sufficientemente diverso perché sia quasi sempre impossibile conformarsi sia alla convenzione che alla sua alternativa. È sufficiente richiedere che nessun segnale di  $\langle Fc, Fa \rangle$  che sia anche segnale di  $\langle Fc', Fa' \rangle$  possa essere vero in entrambi. È più che sufficiente richiedere che i due sistemi non abbiano segnali in comune; in effetti, può essere impossibile, perché può darsi che in S non vi siano abbastanza segnali diversi disponibili. Se i due sistemi di segnalazione hanno in comune alcuni o tutti i segnali, allora, come abbiamo previsto, la veracità sistematica in uno coinciderà a volte o sempre con una certa specie di non veracità sistematica nell'altro.

Quella che *non* ha un'alternativa, e non è una convenzione, è la regolarità seguente: in qualunque esempio di S fra membri di P, il comunicante cerca di emettere qualunque segnale sia vero nella *convenzione prevalente* in quell'esempio, e l'uditorio risponde facendo tutto ciò che sembra meglio nell'ipotesi che egli sia riuscito nel suo intento. Questa regolarità vale ogniqualvolta i membri di P hanno un sistema di segnalazione convenzionale per S con segnali indicativi, quale che possa essere la loro convenzione. Possiamo dire che un segnale  $\sigma$  è *vero* in P in un esempio di S se e solo se c'è qualche

opportuno sistema di segnalazione che sia adottato convenzionalmente in P e  $\sigma$  è vero in quel sistema di segnalazione in quell'esempio di S.

I segnali imperativi possono essere trattati in modo analogo. Sia  $\sigma$  un segnale imperativo di fare r in un sistema di segnalazione  $\langle Fc, Fa \rangle$  per un problema di segnalazione S. Possiamo chiamare  $\sigma$  *vero* in  $\langle Fc, Fa \rangle$  in un esempio di S se quasi tutti i membri dell'uditorio fanno r in quell'esempio, e altrimenti *falso* in  $\langle Fc, Fa \rangle$  in quell'esempio di S.

Perché “vero”? Perché non “eseguito”? Ammetto che il mio uso di “vero” non sia quello corrente; ma “eseguito” non sarebbe affatto quello che intendo. Un imperativo condizionale, “Fa’ così-e-così nel caso così-e-così”, è eseguito se l’antecedente si rivela falso? Forse no, ma io intendo dire che è vero. Un imperativo è forse eseguito se l’imperativo stesso non fa parte dei motivi che l’esecutore ha di fare ciò che gli è stato ordinato? Probabilmente no, ma io intendo dire che è vero. Per svolgere il ruolo che io assegno a “vero”, Nicolas Rescher introduce una nuova espressione: “portato a termine” (*terminated*);<sup>2</sup> chi non sia disposto a tollerare il mio abuso di “vero” può sostituirlo con “vero o portato a termine”. Ma la mia terminologia ha il vantaggio di mettere in evidenza una certa simmetria tra gli indicativi e gli imperativi: il segnale asserisce o comanda il valere di un certo stato di cose, e è vero se quello stato di cose vale effettivamente.

Supponiamo che  $\langle Fc, Fa \rangle$  sia il sistema di segnalazione di qualche popolazione. Possiamo allora descrivere così la regolarità del suo comportamento: in qualunque esempio di S, ogni membro dell'uditorio cerca di rispondere in modo che il segnale del comunicante sia vero in  $\langle Fc, Fa \rangle$  in quell'esempio, e il comunicante emette qualunque segnale gli sembri più opportuno nell'ipotesi che l'uditorio riesca nel suo intento. Ancora una volta, possiamo descrivere la loro convenzione come una convenzione di veracità in  $\langle Fc, Fa \rangle$ . Ma nel caso indicativo spetta al comunicante badare a che il segnale sia vero, scegliendo per l'emissione il segnale giusto; mentre nel caso imperativo spetta all'uditorio rendere vero il segnale del comunicante rispondendo correttamente a esso. Di nuovo, l'alternativa a una convenzione di veracità in  $\langle Fc, Fa \rangle$  sarebbe una regolarità di veracità in qualche altro sistema di segnalazione  $\langle Fc', Fa' \rangle$  per lo stesso problema di segnalazione, con  $\langle Fc', Fa' \rangle$  sufficientemente diverso da  $\langle Fc, Fa \rangle$ .

Se  $\langle Fc, Fa \rangle$  è un qualunque sistema di segnalazione con segnali indicativi o imperativi, possiamo identificare  $\langle Fc, Fa \rangle$  specificando tre cose. (1) Dobbiamo specificare l'insieme dei segnali del sistema; cioè l'insieme che deve essere codominio di Fc e dominio di Fa. (2) Dobbiamo specificare,

---

<sup>2</sup>*The Logic of Commands*, Dover, New York, 1966, pp. 52-61.

per ciascun segnale del sistema, se è indicativo o imperativo. (3) Dobbiamo specificare, per ciascun segnale del sistema, lo stato di cose in cui è vero. Possiamo assumere che questo stato di cose sia un certo insieme di esempi possibili di S: quelli in cui, come si suoi dire, lo stato di cose vale. In altre parole, specifichiamo le condizioni di verità per i segnali del sistema.

Date queste tre informazioni, Fc e Fa sono determinati. Osservando le condizioni di verità di tutti i segnali del sistema troviamo tutti gli esempi possibili del problema di segnalazione cui il sistema si applica. Se i segnali sono specificati come indicativi, osservando le condizioni di verità dei segnali ricostruiamo Fe; se i segnali sono specificati come imperativi, osservando le condizioni di verità dei segnali ricostruiamo Fa. Dato Fc o Fa, troviamo l'altro semplicemente cercando un piano circostanziale discrezionale che combinandosi con il piano dato dia quella dipendenza delle risposte da stati di cose che viene preferita.

È possibile specificare  $\langle Fc, Fa \rangle$  per mezzo di un'unica funzione  $\mathcal{L}$  costruita nel modo seguente. Il dominio di  $\mathcal{L}$  dev'essere l'insieme dei segnali del sistema. Dato un qualunque segnale  $\sigma$  del sistema,  $\mathcal{L}$  deve assegnargli una *interpretazione*  $\langle \mu, \tau \rangle$ . Il componente  $\mu$ , detto *modo*, indica se  $\sigma$  è indicativo o imperativo. Non importa che cosa esattamente sia  $\mu$ . Assumiamo che sia un numero di codice: 0 per l'indicativo, 1 per l'imperativo. Il componente  $\tau$  dell'interpretazione, detto *condizione di verità*, indica lo stato di cose in cui  $\sigma$  è vero. Possiamo assumere che  $\tau$  sia un insieme di esempi possibili di un qualche problema di segnalazione, cioè di quegli esempi in cui  $\sigma$  è vero nel senso peculiare al suo modo.

Se  $\langle Fc, Fa \rangle$  è un sistema di segnalazione verbale, i suoi segnali sono azioni consistenti nell'emissione di espressioni verbali. Possiamo quindi introdurre una lieve modificazione. Sia il dominio di  $\mathcal{L}$  l'insieme delle espressioni verbali emesse nei segnali del sistema, anziché l'insieme dei segnali stessi. Possiamo chiamare una qualunque espressione verbale del dominio di  $\mathcal{L}$  *enunciato* di  $\mathcal{L}$ . Dopo aver introdotto questa modificazione, è naturale chiamare *lingua* la funzione  $\mathcal{L}$ : lingua associata a  $\langle Fc, Fa \rangle$ . Non che ogni lingua sia associata in tal modo a un sistema di segnalazione verbale. Ma alcune lingue rudimentali lo sono, e sono queste che abbiamo ora esaminato abbastanza dettagliatamente. Rimane da vedere come queste lingue di segnalazione verbale non siano adeguate al modello di lingue più interessanti.

Un enunciato  $\sigma$  di una lingua di segnalazione verbale  $\mathcal{L}$  è *vero* in  $\mathcal{L}$  in un particolare esempio del problema di segnalazione cui  $\mathcal{L}$  si applica se e solo se quell'esempio appartiene alla condizione di verità assegnata a  $\sigma$  da  $\mathcal{L}$ ; altrimenti  $\sigma$  è falso in  $\mathcal{L}$  in quell'esempio del problema. Supponiamo che il sistema di segnalazione associato sia adottato convenzionalmente in

una qualche popolazione; allora possiamo dire che quella popolazione ha la convenzione di usare la lingua  $\mathcal{L}$ . Tale convenzione può essere riformulata così: in qualunque esempio del problema di segnalazione cui  $\mathcal{L}$  si applica, una parte - i comunicanti per gli indicativi, l'uditorio per gli imperativi - cerca di assicurare che il comunicante emetta un enunciato che è vero in  $\mathcal{L}$  in quell'esempio; l'altra parte agisce nel modo che sembra migliore nell'ipotesi che la prima parte sia riuscita nel suo intento. Questa è la convenzione di segnalazione familiare, riformulata adesso come convenzione di veracità in  $\mathcal{L}$ ; le sue alternative sarebbero le regolarità di veracità in altre lingue di segnalazione verbale che si applicano allo stesso problema di  $\mathcal{L}$  ma che sono sufficientemente diverse da  $\mathcal{L}$ .

## 4.5 Significato<sub>nn</sub> convenzionale dei segnali

H. P. Grice, nel suo saggio "Meaning", attira la nostra attenzione sull'evidente differenza illustrata nelle seguenti coppie.<sup>3</sup>

- (1A) Erode dona a Salomè la testa di san Giovanni Battista posta su un vassoio.
- (1B) Erode dice a Salomè: "È morto".
- (2A) Sentendosi svenire, un bambino fa vedere a sua madre quanto è pallido (sperando che ella possa trarre le sue conclusioni e aiutarlo).
- (2B) Un bambino dice a sua madre: "Mi sento svenire."
- (3A) Lascio in mostra il vaso di porcellana che mia figlia ha rotto perché mia moglie lo veda.
- (3B) Dico a mia moglie: "Nostra figlia ha rotto il vaso di porcellana."
- (4A) Il bigliettaio dell'autobus bussa sul pannello che sta dietro al conducente perché questi si giri e veda che l'autobus è al completo.
- (4B) Il bigliettaio dell'autobus suona tre volte il campanello per informare il conducente che l'autobus è al completo.
- (5A) Un poliziotto ferma una macchina mettendosi sulla sua strada.
- (5B) Un poliziotto ferma una macchina facendole segno.

La differenza che ricerchiamo deve essere sottile. I casi A e B sono simili sotto molti aspetti. Possiamo confrontarli con questi casi, per esempio:

---

<sup>3</sup> *Philosophical Review*, 66 (1957), pp. 377-388.

- (1C) Erode lascia la testa da qualche parte; a Salomè capita di vederla.
- (1D) Erode lascia la testa dove sa che Salomè la vedrà, supponendo correttamente che ella non capirà che l'ha lasciata lì perché lei la veda.
- (1E) Erode lascia la testa dove sa che Salomè la vedrà, supponendo erroneamente che ella non capirà che l'ha lasciata lì perché lei la veda.

In tutti i casi A e B, e anche in (1C), (1D) e (1E), è soddisfatta questa prima condizione:

Qualcuno fa una qualche azione che produce una credenza o qualche altra risposta in un uditorio.

In tutti i casi A e B, e anche in (1D) e (1E), ma non in (1C), è soddisfatta questa seconda condizione:

Egli intende - si aspetta e vuole - produrre quella risposta con la sua azione.

In tutti i casi A e B, e anche in (1E), ma non in (1C) o (1D), è soddisfatta questa terza condizione:

L'uditorio si rende conto della sua intenzione di produrre quella risposta con la sua azione.

In tutti i casi A e B, ma non in (1C), (1D) e (1E), è soddisfatta questa quarta condizione:

Egli intende che l'uditorio si renda conto della sua intenzione di produrre quella risposta con la sua azione.

La differenza che Grice ha scoperto tra i casi B e i casi A è data da una quinta condizione:

Egli intende che il riconoscimento, da parte dell'uditorio, della sua intenzione di produrre quella risposta sia efficace ai fini della produzione della risposta. Egli non ritiene che sia una conclusione scontata il fatto che la sua azione produca la risposta voluta, che la sua intenzione venga riconosciuta o meno.

Questa quinta condizione vale in ogni caso B, ma non nei casi A e nei casi come (1C), (1D) e (1E).

Ogniqualevolta vale la quinta condizione, Grice dice che l'agente *significa<sub>nn</sub>* qualcosa con la sua azione. Più concisamente, qualcuno *significa<sub>nn</sub>* qualcosa con un'azione se e solo se intende che l'azione "produca qualche effetto in un uditorio per mezzo del riconoscimento della sua intenzione [di produrre l'effetto, cioè]." <sup>4</sup> "Significa<sub>nn</sub>" è un'abbreviazione di "signifi-

---

<sup>4</sup>"Meaning", p. 385.

ca non naturalmente”: non per suggerire alcunché di soprannaturale, ma in contrapposizione con il cosiddetto significato naturale dei segni naturali: macchie che significano morbillo, fumo che significa fuoco, e simili.

Io avrei sottolineato un'altra differenza tra casi A e casi B. Nei casi B, ma non in quelli A, la risposta dell'uditorio è pro-dotta per mezzo di un segnale convenzionale, dato in conformità a una convenzione di segnalazione. Ma non sono in disaccordo con Grice, in quanto il significato<sub>nm</sub> è una conseguenza della segnalazione convenzionale.

Sia  $\sigma$  un segnale di un sistema di segnalazione convenzionale  $\langle Fc, Fa \rangle$ . Ci siano uno stato di cose  $s$  e una risposta  $r$  tali che  $\sigma$  sia  $Fc(s)$  e  $r$  sia  $Fa(\sigma)$ . Supponiamo che io sia il comunicante e tu l'uditorio in un problema di segnalazione della specie opportuna; e che, avendo osservato che  $s$  vale, io faccia  $\sigma$  in conformità alla nostra convenzione. Supponiamo che si tratti di un caso assolutamente normale: nessuna delle clausole relative alle eccezioni elencate nel capitolo 2.4 si applicano a esso. In tal caso io significo<sub>nm</sub> qualcosa per mezzo di  $\sigma$ . Io intendo che  $\sigma$  produca la tua risposta  $r$  per mezzo del riconoscimento da parte tua di questa intenzione. Dimostrazione:

L'intenzione con cui faccio  $a$  può venir determinata esaminando il ragionamento pratico che mi giustifica nel farlo. Non è necessario che io sviluppi effettivamente il ragionamento per avere un'intenzione; le azioni compiute senza rifletterci sono spesso compiute con intenzioni definite. Esaminiamo quindi la mia giustificazione, in una delle versioni considerate nel paragrafo 2. Essa era illustrata nella Figura 45 (ripetuta qui sotto).

La mia decisione di fare  $a$ , dopo aver notato  $s$ , ha come premessa la mia aspettativa di potere in quel modo produrre  $r$  e dal mio desiderio di produrre  $r$ . Così io faccio  $\sigma$  con l'intenzione di produrre  $r$ .

Io mi aspetto che tu deduca  $s$  notando che io faccio  $\sigma$ . Mi aspetto che tu riconosca il mio desiderio di produrre  $r$ , subordinatamente a  $s$ . Mi aspetto che tu riconosca la mia aspettativa di poter produrre  $r$  facendo  $\sigma$ . Così mi aspetto che tu riconosca la mia intenzione di produrre  $r$ , quando noti che faccio  $\sigma$ .

Anche la mia aspettativa di poter produrre  $r$  facendo  $\sigma$  ha come premessa la mia aspettativa che tu riconosca il mio desiderio di produrre  $r$ , subordinatamente a  $s$ , e la mia aspettativa che tu riconosca la mia aspettativa di poter produrre  $r$  facendo  $\sigma$ . Queste due premesse costituiscono sia la mia aspettativa che tu riconosca la mia intenzione, sia il mio motivo di aspettarmi che

la mia intenzione venga esaudita. Così io mi aspetto che il riconoscimento da parte tua della mia intenzione sia efficace ai fini della produzione della tua risposta. Non ritengo che sia una conclusione scontata il fatto che la mia azione produca  $r$ , che tu riconosca o meno la mia intenzione di produrre  $r$ .

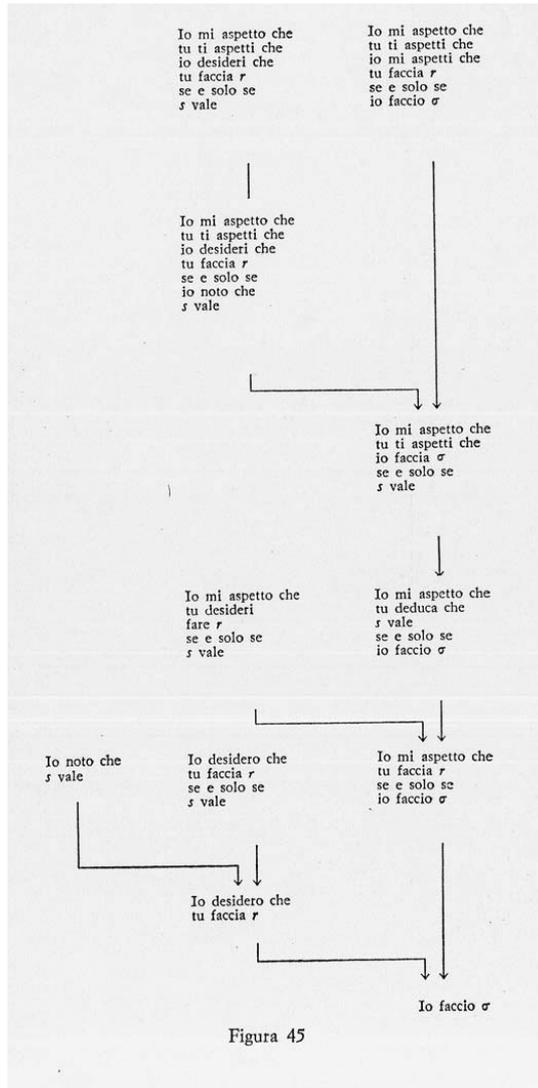


Figura 45

Un segnale convenzionale, emesso in assoluta conformità alla convenzione opportuna, deve significare. qualcosa. È vera la converso? No, per diversi

motivi.

Può darsi che io vi stia ingannando. Il mio segnale significherebbe<sub>nn</sub> qualcosa, come al solito, ma io violerei la convenzione.

Oppure potrei sbagliarmi riguardo alle convenzioni prevalenti e significare<sub>nn</sub> qualcosa per mezzo di un'azione che ho creduto erroneamente che fosse un segnale convenzionale. Il mio ragionamento, le mie aspettative, le mie preferenze, e quindi le mie intenzioni, sembreranno gli stessi sia che io sia nel giusto sia che mi sbagli. L'unica differenza consisterà nel fatto che non riuscirò a produrre la risposta voluta.

Oppure potrei aspettarmi (correttamente o erroneamente) che il mio uditorio si sbagli riguardo alle convenzioni prevalenti. Questo è quanto accade nell'esempio seguente, tratto da "What is a Speech Act?" di John Searle.

"Supponiamo che io sia un soldato americano durante la seconda guerra mondiale e che sia catturato dalle truppe italiane. Supponiamo anche che, per farmi liberare, io voglia far credere a queste truppe che sono un ufficiale tedesco. Quello che vorrei fare è dir loro in tedesco o in italiano che sono un ufficiale tedesco. Ma supponiamo che io non conosca abbastanza l'italiano o il tedesco per farlo. Perciò tento, per così dire, di recitare la scena della rivelazione della mia identità di ufficiale tedesco facendo uso di quel po' di tedesco che conosco, nella speranza che essi non sappiano il tedesco tanto da scoprire il mio piano. Supponiamo che io conosca soltanto una riga di tedesco, che ricordo da una poesia che ho dovuto imparare a memoria durante un corso di tedesco al liceo. Quindi io, un prigioniero americano, mi rivolgo con questa frase a coloro che mi hanno catturato: "Kennst du das Land, wo die Zitronen blühen?" Io intendo produrre [...] come effetto la credenza che io sia un ufficiale tedesco; e intendo produrre questo effetto per mezzo del riconoscimento da parte loro della mia intenzione."<sup>5</sup>

Non è essenziale che il soldato cerchi di ingannare chi l'ha catturato. Potrebbe altrettanto bene essere un vero ufficiale tedesco che ha dimenticato il resto del suo tedesco in seguito al trauma provocato dalla guerra.

Oppure potrei significare<sub>nn</sub> qualcosa senza pensare che mi stavo conformando a una convenzione, e senza aspettarmi che neppure il mio uditorio lo pensi. P.F. Strawson dà un esempio di questo genere in "Intention and Convention in Speech Acts".

---

<sup>5</sup> *Philosophy in America*, a cura di Max Black, Cornell University Press, Ithaca, 1965, pp. 229-230.

“S [...] prepara una “prova” apparentemente convincente che p, in un luogo in cui A non può non vederla. Egli fa questo sapendo che A sta osservandolo mentre è all’opera, ma sapendo anche che A non sa che S sa che A lo osserva mentre è all’opera. Egli capisce che A non considererà la prova predisposta come una prova autentica o naturale che p, ma capisce, anzi vuole, che A consideri la sua pre-disposizione della prova come un motivo per pensare che egli, S, intende produrre in A la credenza che p [...] Egli sa che A ha dei motivi generali per pensare che S non vorrebbe indurre lui, A, a pensare che p se non sapesse che effettivamente si dà il caso che p; e sa quindi che il riconoscimento da parte di A della sua (di S) intenzione di indurre A a credere che p sembrerà effettivamente a A un motivo sufficiente per credere che p. Ed egli vuole che il riconoscimento da parte di A della sua intenzione [...] funzioni proprio in questo modo.”<sup>6</sup>

In questo caso, se facessimo regredire di un altro passo la giustificazione che S dà della propria azione, emergerebbe una anomalia. *Non* è vero che S si aspetta che A si aspetti che S si aspetti che A si aspetti che S predisponga la “prova” se sa che p.

Oppure potrei significare<sub>nn</sub> qualcosa senza pensare che mi stavo conformando a una convenzione, senza aspettarmi che neppure il mio uditorio lo pensi, e con esattamente la stessa specie di giustificazione per ricostruzione che avrei se mi conformassi a una proposizione. Supponiamo che io abbia scoperto delle sabbie mobili e non conosca alcun segnale di allarme convenzionale. Immergo fino al petto un pupazzo nelle sabbie mobili, sperando che chi lo vede capisca. Non esiste la convenzione di contrassegnare le sabbie mobili con pupazzi semi-immersi; né io penso che esista, né mi aspetto che il mio uditorio lo pensi. Ma voglio che la mia azione produca la consapevolezza che qui ci sono le sabbie mobili per mezzo del riconoscimento della mia intenzione di produrre quella consapevolezza. Ho svolto la mia parte in un sistema di segnalazione all’interno di un problema di segnalazione; e spero che il mio uditorio futuro farà la propria parte. In questo caso, però, la coordinazione sarà raggiunta non in virtù del precedente, ma in virtù della salienza.

Infine, potrei significare<sub>nn</sub> qualcosa per mezzo di un’azione in conformità a una convenzione di linguaggio, ma non in conformità a una convenzione di segnalazione. Abbiamo trovato plausibile che alcuni usi del linguaggio, ma non tutti, siano regolati da convenzioni di segnalazione verbali, oltre che

---

<sup>6</sup> *Philosophical Review*, 73 (1964), pp. 446-447.

da convenzioni di linguaggio più generali. Anche dopo aver esaminato quale potrebbe essere la configurazione delle convenzioni generali di linguaggio, non cercherò di dimostrare che un segnale emesso in conformità a esse deve significare<sub>nn</sub> qualcosa; ma avanzo l'ipotesi che sia così.

Searle trae questa morale dall'esempio del soldato: “dobbiamo catturare tanto gli aspetti intenzionali che quelli convenzionali [della comunicazione] e soprattutto le loro relazioni.”<sup>7</sup> Io ho sostenuto che, una volta che abbiamo catturato l'aspetto convenzionale, abbiamo finito. Abbiamo catturato anche l'aspetto intenzionale.

---

<sup>7</sup> “What is a Speech Act?” p. 230.



## Capitolo 5

# Convenzioni di linguaggio

### 5.1 Lingue possibili

Una lingua di segnalazione verbale  $\mathcal{L}$  è stata chiamata, correttamente, una lingua. Ma è una lingua rudimentale, perlomeno per i motivi seguenti.

C'è solo un insieme chiuso, finito, di enunciati di  $\mathcal{L}$ . Le condizioni di verità si possono dare enunciato per enunciato. Non è possibile creare un enunciato nuovo, con la sua condizione di verità, a partire dalle parti di cui già disponiamo.

Non c'è nulla che assomigli a una conversazione frivola. L'uso degli enunciati di  $\mathcal{L}$  è destinato ad alcune attività particolari, con il solo scopo di proseguire con successo quell'attività. Perciò, se è usata in base a una convenzione, tale convenzione è mantenuta da un interesse a breve scadenza per la coordinazione. La mancata conformazione da parte di qualcuno, quando gli altri si conformano, provoca direttamente nell'uditorio una risposta che, dato lo stato di cose attuale, è indesiderata.

Gli utenti di  $\mathcal{L}$  hanno scarsa libertà di scelta. Nel caso indicativo, un comunicante che ha osservato un certo stato di cose e desidera essere verace in  $\mathcal{L}$  deve emettere un certo enunciato. Non ha scelta tra il parlare e lo star zitto; non ha scelta per quanto riguarda l'oggetto di cui parlare; non ha scelta neppure per quanto riguarda il modo di formulare il proprio messaggio.

Per qualunque enunciato indicativo di  $\mathcal{L}$ , abbiamo stipulato che il comunicante sia in condizione di dire se è vero in  $\mathcal{L}$  in una occasione particolare. Non c'è posto per enunciati indicativi che esprimono opinioni personali o ipotesi provvisorie.

Un enunciato indicativo di  $\mathcal{L}$  è vero in un esempio di una certa situazione se in quell'esempio vale un certo stato di cose. Gli enunciati indicativi

non esprimono mai fatti generali, ma soltanto fatti relativi all'occasione di emissione dell'enunciato. (Per i motivi finora elencati, gli enunciati indicativi di  $\mathcal{L}$  sono inadeguati a esprimere la speculazione, la decisione o l'argomentazione.)

L'uditorio è in condizione di render vero un qualunque enunciato imperativo di  $\mathcal{L}$  per mezzo di una risposta nella situazione in cui viene emesso. Non ci sono enunciati imperativi che comunichino consigli di carattere generale o disposizioni permanenti.

L'uditorio ha interesse a render vero qualunque enunciato imperativo di  $\mathcal{L}$  emesso da un comunicante che si aspetta che sia reso vero. Non c'è bisogno di decidere se il comunicante ha le conoscenze e il giudizio necessari per dare un consiglio valido, o se ha l'autorità di dare ordini, o il potere e la volontà di imporre le proprie richieste, e così via.

Non ci sono abbastanza modi. Sebbene sia facile moltiplicare i modi oltre il necessario, non c'è una maniera plausibile di procedere con i nostri due soltanto. Forse i modi addizionali più urgenti sono: l'interrogativo; il commissivo, per le promesse e le minacce; il permissivo, per l'esplicita astensione dagli imperativi.<sup>1</sup> Può darsi che alcuni di questi siano riducibili agli indicativi e agli imperativi, ma tutte le riduzioni di questo genere sono, nel migliore dei casi, problematiche.

In  $\mathcal{L}$  non c'è ambiguità o indicialità. Ciascun enunciato ha un unico modo fisso e un'unica condizione di verità fissa, la stessa in ogni possibile occasione di emissione in  $\mathcal{L}$ .

Le lingue di segnalazione verbale possono essere difettose anche in altri modi. Ma ne abbiamo visti a sufficienza. Passiamo a una classe più ampia di lingue possibili. Spero che sarà ampia abbastanza da comprendere lingue come il gallese, o l'inglese, o l'esperanto. Se no, spero che comprenderà almeno larghi frammenti centrali di queste lingue.

In questo paragrafo considereremo che cos'è una lingua *possibile*, facendo astrazione dagli utenti che essa può avere e anche dal problema di come sarebbe usata se avesse degli utenti. In seguito considereremo che cosa significa per una popolazione usare una lingua - in altre parole, che cosa fa di una lingua *possibile* la lingua *attuale* di qualcuno.

Con "lingua possibile" intendo semplicemente un'entità della specie che sto per specificare. Non voglio dire che ogni entità di questa specie è se-

---

<sup>1</sup>Cfr. R.M. Hare, "Some Alleged Differences between Imperatives and Indicatives", *Mind*, 76 (1967), pp. 309-326, per una spiegazione dei permissivi come imperativi esplicitamente trattenuti. Hare suggerisce che si possa distinguere un modo analogamente collegato al modo indicativo: "un modo loquace di non fare una certa asserzione" (p. 321).

riamente candidata ad essere la lingua attuale di una popolazione umana. Ci saranno lingue possibili il cui uso non potrebbe servire ad alcun intento dell'uomo, lingue possibili così ingombranti che non potrebbero mai venir adottate, e addirittura lingue possibili che gli uomini sono psicologicamente o fisiologicamente incapaci di usare.

Parlando di una lingua, intendo riferirmi a una lingua *interpretata*. Così quella che io chiamo lingua è ciò che molti logici chiamerebbero lingua più un'interpretazione per essa. Secondo il mio punto di vista non ci possono essere due interpretazioni della stessa lingua; ma ci possono essere due lingue con gli stessi enunciati.

Partiremo dalle lingue possibili che abbiamo già visto: le lingue di segnalazione verbale. Ricordiamo che una lingua di questo tipo è una funzione  $\mathcal{L}$  che assegna a ogni espressione verbale di un qualche insieme finito - a ogni enunciato di  $\mathcal{L}$  - un'interpretazione consistente di un modo  $\mu$  e di una condizione di verità  $\tau$ . Il modo  $\mu$  è un numero di codice: 0 per gli indicativi, 1 per gli imperativi. La condizione di verità  $\tau$  è un insieme di esempi possibili del problema di segnalazione cui  $\mathcal{L}$  si applica: l'insieme di quegli esempi in cui l'enunciato è vero, nel senso indicativo (se  $\mu$  è 0) o nel senso imperativo (se  $\mu$  è 1). Adesso, ai fini della generalità, dobbiamo perfezionare questa descrizione punto per punto.

In primo luogo, abbiamo visto che ci dovrebbero essere più di due modi. Permetteremo che una lingua possibile abbia un numero finito qualunque di modi, poiché non conosciamo un limite superiore da imporre al numero di cui avremo bisogno. Ma, siccome un limite superiore probabilmente esiste, e probabilmente non è molto alto, otterremo molte cosiddette lingue possibili prive di interesse. Ma questo non importa; quello che importa è non trascurarne di interessanti. Questi nuovi modi possono essere semplicemente degli ulteriori numeri di codice, oltre 1 e 0: per esempio, 2 per gli interrogativi, 3 per i commissivi, 4 per i permissivi, e così via, in un certo ordine, per qualunque altro modo sia necessario. Se, come speriamo, alcuni di questi modi sono riducibili agli indicativi e agli imperativi, possiamo semplicemente lasciare inutilizzati alcuni numeri di codice.

Quali differenze ci sono tra i modi, tali da giustificare il fatto di chiamarli indicativo, imperativo, e così via? Nessuna, per ora. Finché facciamo astrazione dall'uso di una lingua possibile, i modi non sono che numeri che compaiono nelle interpretazioni. Ma in seguito, quando vedremo che cosa ci vuole per fare di una lingua possibile la lingua attuale di una popolazione, i diversi numeri di codice giocheranno dei ruoli diversi nelle regolarità convenzionali per mezzo delle quali la lingua viene usata. Soltanto allora i nomi che abbiamo assegnato ai numeri di codice risulteranno giustificati.

Abbiamo già visto come ciò avvenga nel caso di una lingua di segnalazione verbale: quando la lingua è usata in base a una convenzione, se  $\mu$  è 0 è il comunicante che cerca di render veri i propri enunciati, se  $\mu$  è 1 è l'uditorio che cerca di render veri gli enunciati del comunicante.

In secondo luogo, non possiamo più concepire le condizioni di verità come insiemi di esempi possibili della situazione cui la lingua si applica; non c'è più alcuna situazione particolare cui la lingua si applichi. Possiamo invece concepire una condizione di verità come un insieme di mondi possibili: l'insieme dei mondi in cui, come si suoi dire, vale la condizione di verità. Se  $\mathcal{L}$  assegna un'interpretazione  $\langle \mu, \tau \rangle$  a un enunciato  $\sigma$ , allora  $\tau$  è l'insieme dei mondi in cui  $\sigma$  è vero in un senso appropriato a  $\mu$ .

In terzo luogo,  $\mathcal{L}$  può contenere enunciati indicali le cui condizioni di verità dipendono da colui che li emette, dall'uditorio cui egli li destina, o dal momento e dal luogo dell'emissione.  $\mathcal{L}$  può contenere enunciati anaforici le cui condizioni di verità dipendono dal contesto del discorso precedente o del discorso che ci si propone di fare in seguito. (Cioè enunciati come "Poi egli si tolse il soprabito" o "La parte suddetta si rifiutò di pagare.")  $\mathcal{L}$  può contenere enunciati le cui condizioni di verità dipendono dalle circostanze della loro emissione: "Chiudi la porta" o "Il sale è sulla tavola."  $\mathcal{L}$  deve quindi assegnare interpretazioni non agli enunciati stessi, ma agli enunciati in occasioni possibili della loro emissione. Un'occasione possibile di emissione di un enunciato  $\sigma$  può forse essere identificata con una coppia formata da un mondo possibile e da una posizione spazio-temporale in esso, tale che  $\sigma$  sia emesso in quella posizione in quel mondo. Dati il mondo e la posizione, si può presumere che sia disponibile ogni ulteriore informazione relativa al contesto, di cui abbiamo bisogno. Avremo identificato in modo univoco l'emittente, l'uditorio cui egli si rivolge, il discorso precedente, le circostanze, e così via. (In quei bizzarri mondi possibili in cui in una posizione può verificarsi più di una emissione di un enunciato, sarebbe necessario inserire, in quell'entità che concepiamo come occasione dell'emissione, più informazioni relative al contesto. Ma non c'è alcun bisogno che tutte le occasioni di emissione degli enunciati siano entità dello stesso tipo.)

In quarto luogo, non possiamo assumere che gli enunciati di  $\mathcal{L}$  siano non ambigui.  $\mathcal{L}$  deve quindi assegnare non una interpretazione ma un insieme di interpretazioni. Quando, in un'occasione di emissione, l'enunciato non è ambiguo, l'insieme conterrà esattamente una interpretazione; quando l'enunciato è ambiguo, l'insieme conterrà un numero finito di interpretazioni. Probabilmente è opportuno permettere un terzo caso: l'insieme potrebbe non contenere alcuna interpretazione. Questo è un modo di trattare enunciati come "La porta è aperta" in un'occasione della sua emissione in cui

nessuna porta sia presente o sia stata menzionata; ed enunciati anomali come “La pentola è morta”, “Ogni ragazza si cantò un cavallo” -o “La quadruplicità beve la procrastinazione” in tutte le occasioni della loro emissione. (Un’alternativa consisterebbe nel trattare gli enunciati anomali come non enunciati, a costo di complicare la grammatica della lingua; ma tale trattamento non andrebbe bene per “La porta è aperta”, in quanto non vogliamo che l’essere o meno enunciato dipenda dalle occasioni di emissione. Un’altra alternativa consisterebbe nell’assegnare interpretazioni con condizioni di verità vuote, come faremmo con enunciati autocontraddittori. Quest’ultima potrebbe essere l’alternativa migliore; ma preferisco lasciare la questione aperta.) Potrebbe darsi che vogliamo stipulare che tutte le interpretazioni assegnate a un enunciato in un’occasione della sua emissione abbiano lo stesso modo; ma non dobbiamo farlo. Si consideri “Cerca di chiudere la porta” in un’occasione della sua emissione in cui una porta aperta bene in vista fa infreddolire tutti, e in un’occasione in cui l’emittente ha anche raccontato una storia e in cui gli è appena stato chiesto: “E adesso che cosa fa?”

Abbiamo ormai accumulato un numero notevole di modificazioni da apportare alla funzione  $\mathcal{L}$ . Adesso abbiamo una funzione i cui argomenti sono coppie formate da un enunciato e da un’occasione possibile della sua emissione, e i cui valori sono insiemi di interpretazioni finiti, eventualmente vuoti. Un’interpretazione è ancora una coppia formata da un modo e da una condizione di verità; ma un modo può essere, adesso, un numero naturale qualunque (ma sarà piccolo, se  $\mathcal{L}$  è una lingua interessante) e una condizione di verità è un insieme di mondi possibili.

Inizialmente sarebbe stato ragionevole dire che il significato di un enunciato è dato dalla singola interpretazione assegnatagli da  $\mathcal{L}$ , indipendentemente dalle sue occasioni di emissione. Ma ora non possiamo più dire che il significato - o addirittura un significato - di un enunciato sia dato da una qualunque delle sue interpretazioni. Ciò non basterà neppure per un enunciato che non è ambiguo in nessuna delle occasioni della sua emissione. Nell’interpretazione che esso riceve in una qualunque occasione, infatti, non si manifesterà alcuna delle caratteristiche indicali o anaforiche dell’enunciato. Ciò che dà il significato dell’enunciato - nella misura in cui esso può essere dato senza considerare il significato delle sue parti - è la funzione in virtù della quale il suo insieme di interpretazione dipende dalle caratteristiche delle occasioni di emissione.

Infine, non possiamo più stipulare che l’insieme degli enunciati della lingua - il dominio di  $\mathcal{L}$ - sia finito. Qualunque lingua interessante ha un numero infinito di enunciati. È vero che verrà effettivamente emesso al massimo un numero finito di enunciati di qualunque lingua. In qualunque mondo possi-

bile in cui i limiti umani rimangano quelli che sono, *potrebbe* venir emesso al massimo un numero finito di enunciati. Ma questi insiemi finiti sono privi di interesse, perché non abbiamo alcuna speranza di scoprire quali enunciati contengono. È del tutto ovvio dire che qualunque utente di una lingua ha, riguardo a enunciati che non vengono mai emessi, la stessa competenza che ha riguardo a enunciati che incontra ogni giorno. Ogniqualvolta ciò avvenga, la lingua dell'utente dovrebbe essere concepita come comprendente tutti gli enunciati che egli *potrebbe* usare e capire se avesse abbastanza pazienza, tempo e memoria. Se non permettiamo che le lingue abbiano un numero infinito di enunciati, dovremo imporre alla lunghezza degli enunciati limiti superiori arbitrari e ingiustificabili.

## 5.2 Grammatiche

Non è vero, però, che qualunque insieme infinito arbitrario di espressioni verbali vada bene come insieme degli enunciati di una lingua interessante. Nessuna lingua adeguata agli scopi dei suoi utenti può essere finita; ma qualunque lingua che debba venir usata da esseri umani finiti deve possedere la caratteristica ottimale immediatamente seguente: quella di essere finitamente specificabile. Deve avere una grammatica finita, in modo che tutti i suoi enunciati, con le loro interpretazioni, possano venir specificati facendo riferimento ai loro costituenti elementari, in numero finito, e alle operazioni, pure in numero finito, per la costruzione di costituenti più ampi a partire da quelli più piccoli.

Buona parte degli sforzi recenti della teoria linguistica sono stati dedicati a ricercare una forma normale opportuna per le grammatiche.<sup>2</sup> Il progetto è di ridurre la classe delle lingue possibili riducendo la classe delle grammatiche possibili, finché le sole lingue possibili rimanenti siano quelle seriamente candidate ad essere usate dall'uomo. In pratica, ciò viene effettuato restringendo al massimo la forma normale delle grammatiche, compatibilmente con l'esigenza di non lasciare priva di grammatica nessuna delle lingue attuali.

Non ci recherà alcun danno il fatto di dover annoverare tra le lingue possibili, oltre a quelle che vogliamo realmente, molte entità extra. Per questo non stipuleremo che le nostre lingue possibili debbano avere grammatiche di una qualunque forma specificata. In realtà, non abbiamo bisogno di inserire il possesso di una qualunque grammatica tra le condizioni definitorie

---

<sup>2</sup>Cfr., per esempio, Noam Chomsky, *Aspects of the Theory of Syntax*, MIT Press, Cambridge, Mass., 1965, Cap. I. [Trad. it. di A. Woolf De Benedetti, *Aspetti della teoria della sintassi*, in *Saggi linguistici*, Boringhieri, Torino, 1970, vol. II (N.d.T.)]

delle lingue possibili. Possiamo risparmiarci la fatica di cercare di dire, con precisione e generalità adeguate, che cosa significa avere una grammatica. Ma dobbiamo ricordarci che le lingue prive di grammatica - o prive di una grammatica avente quella che risulterà essere la forma normale opportuna - si chiamano lingue possibili solo perché siamo stati troppo pigri per escluderle.

Cercherò tuttavia di specificare come potrebbe funzionare una specie di grammatica per una lingua possibile  $\mathcal{L}$ . Distinguerò, in una grammatica, tre parti, chiamate lessico, componente generativo e componente rappresentativo.

Il *lessico* è un ampio insieme finito di costituenti elementari, le cui categorie vengono indicate per mezzo di indicatori. Tali costituenti saranno per la maggior parte parole, o morfemi più brevi delle parole.

Il *componente generativo* è un insieme finito di *operazioni di combinazione*. Ciascuna di queste opera su un certo numero di costituenti appartenenti a categorie date concatenandoli in modo da costruire un nuovo costituente più ampio, appartenente a una data categoria. Essa provvede inoltre questo nuovo costituente di un indicatore, il quale mostra la sua categoria e i costituenti con i quali è costruito. Partendo dagli elementi lessicali, il componente generativo costruisce costituenti sempre più ampi. Più precisamente: un costituente è un membro qualunque del più piccolo insieme contenente gli elementi lessicali e chiuso rispetto alle operazioni di combinazione. I costituenti sono quindi stringhe di elementi lessicali associate a una gerarchia di indicatori di categoria, come viene illustrato nella Figura 46.

Il *componente rappresentativo* opera su alcuni dei costituenti costruiti dal componente generativo - quelli appartenenti alla categoria *enunciato* - per produrre espressioni verbali. Le espressioni verbali che vengono così a rappresentare i costituenti d'enunciato sono gli enunciati di  $\mathcal{L}$ . Nel caso particolare di una *grammatica a struttura sintagmatica*, il componente rappresentativo ha un ruolo assai limitato. Esso toglie semplicemente gli indicatori di categoria e rimpiazza gli elementi lessicali, in ordine, con opportune stringhe di suoni o di caratteri. Le grammatiche delle lingue formalizzate perlomeno di quelle con sistemi di interpunzione semplici - sono grammatiche a struttura sintagmatica. Nel caso più generale di una *grammatica trasformazionale*, il componente rappresentativo ha un ruolo molto più importante. Impiegando le informazioni contenute negli indicatori di categoria, prima di produrre l'espressione verbale, esso può permutare delle parti del costituente d'enunciato, cancellarle o aggiungerne di nuove.

Non stipuleremo che ci debba essere una corrispondenza biunivoca tra costituenti d'enunciato ed espressioni verbali che li rappresentano. Un enun-

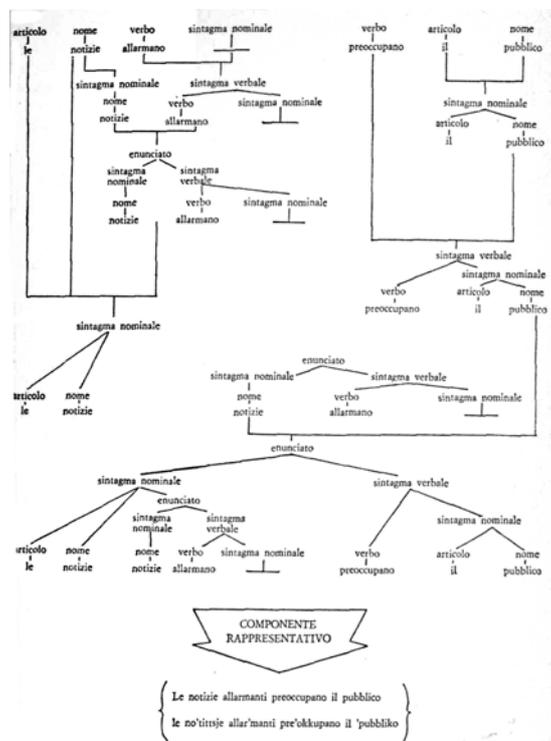


Figura 46

ciato potrebbe rappresentare più costituenti d'enunciato diversi: ambiguità sintattica. Oppure un costituente d'enunciato potrebbe essere rappresentato da più enunciati diversi, ciò che rappresenta un tipo di variazione stilistica. Oppure un costituente d'enunciato potrebbe non venir rappresentato; permettendo che il componente rappresentativo sia selettivo possiamo semplificare il componente generativo.

La grammatica dovrebbe fornire non solo gli enunciati di  $\mathcal{L}$  ma anche la loro interpretazione nelle loro occasioni di emissione. Possiamo conseguire questo scopo (1) assegnando delle interpretazioni vedremo presto che specie di cose sono agli elementi lessicali, (2) fornendo, per ogni operazione di combinazione usata per costruire un nuovo costituente  $\xi$  a partire da costituenti precedenti  $\xi_1 \dots \xi_k$ , una *operazione di proiezione* di accompagnamento per derivare, data un'interpretazione per ciascun  $\xi_1 \dots \xi_k$ , un'interpretazione per  $\xi$ , e (3) trasmettendo le interpretazioni dei costituenti d'enunciato agli enunciati che li rappresentano. Siccome le operazioni di combinazione co-

struiscono un numero infinito di costituenti sempre più grandi, partendo dagli elementi lessicali, le corrispondenti operazioni di proiezione procederanno in parallelo per derivare le interpretazioni di quei costituenti, partendo dalle interpretazioni degli elementi lessicali.

Un enunciato può avere più interpretazioni, le quali gli sono state trasmesse; ciò può accadere in uno e/o nell'altro dei due modi seguenti. L'enunciato può essere ambiguo sintatticamente, in quanto rappresenta - e riceve interpretazioni da - più di un costituente d'enunciato. Un esempio di Chomsky è: "John was frightened by the new methods", che è ambiguo sebbene non contenga parole ambigue.<sup>3</sup> Oppure l'enunciato può essere ambiguo perché rappresenta un costituente d'enunciato che ha già più interpretazioni. Un elemento lessicale può infatti essere ambiguo. Quando viene inserito in un costituente più ampio, le diverse interpretazioni dell'elemento lessicale produrranno in generale diverse interpretazioni del costituente più ampio, E così via, fino a un costituente d'enunciato ambiguo, rappresentato da un enunciato ambiguo. Un esempio è "A Paolo piace la pesca".

Non è necessario assumere che, quando un nuovo costituente  $\xi$  è costruito a partire da costituenti precedenti  $\xi_1 \dots \xi_k$ , *ogni* combinazione possibile di interpretazioni di  $\xi_1 \dots \xi_k$  produca un'interpretazione di  $\xi$ . Se fosse così, l'ambiguità non avrebbe più limiti. Le operazioni di proiezione possono essere selettive, e agire soltanto sulle combinazioni in ingresso di interpretazioni che soddisfano certe restrizioni. Ma in questo caso potrebbe accadere che non riceva alcuna interpretazione, per quanto sia stato costruito adeguatamente a partire da  $\xi_1 \dots \xi_k$ , i quali ammettono tutti delle interpretazioni. Infatti, nessuna combinazione di interpretazioni di  $\xi_1 \dots \xi_k$  potrebbe essere un ingresso accettabile per l'operazione di proiezione. Se  $\xi$ , a sua volta, entra a far parte di un altro costituente, anche quest'ultimo, di norma, non avrà alcuna interpretazione; infatti all'operazione di proiezione non sarà fornita neppure una combinazione completa di interpretazioni sui cui agire. E così via, fino ai costituenti enunciativi privi di interpretazione. Se un simile costituente anomalo d'enunciato viene rappresentato per mezzo di un enunciato, non avrà alcuna interpretazione da trasmettere all'enunciato. È per questo che è possibile che gli enunciati stessi non abbiano alcuna interpretazione. (Non sto sostenendo che qualunque buona grammatica di qualunque lingua familiare debba produrre enunciati privi di interpretazione; ma, come ho già detto, sembra opportuno lasciare aperta questa possibilità.)

Un elemento lessicale può essere indicale, ricevendo interpretazioni di-

---

<sup>3</sup>Un esempio analogo in italiano porrebbe essere: "La vista della linee è impressionante." (N.d.T.)

verse in diverse occasioni possibili della sua emissione. (Diremo che un'occasione di emissione di un costituente è una qualunque occasione di emissione di un enunciato che rappresenta un costituente d'enunciato che lo contiene.) Questa dipendenza dall'occasione si trasmette, attraverso le operazioni di proiezione, a costituenti sempre più ampi, e infine agli enunciati.

Il mio modo di delineare la natura di una grammatica per  $\mathcal{L}$  era destinato ad avere abbastanza generalità da comprendere due casi particolari: (1) gli insiemi delle regole di formazione e di valutazione impiegate dai logici per specificare le lingue formalizzate che studiano, e (2) le grammatiche trasformazionali delle lingue naturali, nella forma recentemente proposta da Chomsky, Jerrold Katz e altri.<sup>4</sup> Esistono proposte anteriori di Chomsky e dei suoi collaboratori le quali permetterebbero grammatiche che non si accordano alla mia descrizione; ma non è stata data alcuna dimostrazione convincente della necessità di questa iniziale generalità. In ogni modo, non importa se ho dato una definizione adeguata di grammatica, purché abbia mostrato, grosso modo, a che cosa assomiglia.

Ho aderito alla spiegazione data da Katz del modo in cui una grammatica deriva l'interpretazione di un enunciato partendo dalle interpretazioni degli elementi lessicali, impiegando le operazioni di proiezione per derivare le interpretazioni di costituenti sempre più ampi, e infine trasmettendo le interpretazioni dai costituenti d'enunciato agli enunciati che li rappresentano. Ma non ho accettato la spiegazione data da Katz della natura di tali interpretazioni; questo è un problema a sé. Katz le concepisce come espressioni costruite con simboli chiamati "indicatori semantici", che rappresentano "gli elementi concettuali della struttura di un senso".<sup>5</sup> Trovo insoddisfacente questa spiegazione, in quanto ha come sbocco una teoria semantica che esclude nozioni semantiche centrali come quelle di verità e di riferimento.

Che cos'è, allora, l'interpretazione di un costituente? Un caso l'abbiamo già risolto. L'interpretazione di un enunciato, e quindi di un costituente d'enunciato, è una coppia formata da un modo e da una condizione di verità: un numero di codice e un insieme di mondi possibili. Il modo è qualcosa di peculiare agli enunciati; mentre la condizione di verità suggerisce una strategia generale per assegnare ai costituenti interpretazioni appropriate.

La semantica referenziale di ispirazione tarskiana e carnapiana assegna

---

<sup>4</sup>Chomsky, *Aspects of the Theory of Syntax*, cit.; Jerrold Katz, *Philosophy of Language*, Harper and Row, New York, 1966, cap. IV. I miei costituenti, in quanto sono associati a gerarchie di indicatori di categoria, coincidono con gli indicatori sintagmatici sottostanti di Chomsky, o con sottoalberi di essi.

<sup>5</sup>Katz, *Philosophy of Language*, pp. 155-156.

ai costituenti interpretazioni adatte alle categorie cui appartengono: valori di verità per gli enunciati, denotazioni per i nomi, insiemi per i predicati a un argomento, insiemi di n-uple per i predicati a n argomenti, ecc. Date le estensioni degli elementi lessicali, le opportune estensioni di costituenti sempre più ampi sono derivabili per mezzo delle operazioni di proiezione (regole di valutazione). Sarebbe bene, naturalmente, che le interpretazioni non fossero semplici estensioni, poiché l'estensione dipende sia dall'interpretazione che da fatti accidentali relativi al mondo attuale; per esempio, costituenti che dovrebbero avere interpretazioni diverse risultano accidentalmente coestensivi. Ciononostante, la semantica referenziale sembra fallire in parte il bersaglio.

Una condizione di verità specifica i valori di verità di un enunciato; ma in tutti i mondi possibili, non solo in quello che si trova a essere attuale. Possiamo interpretare un costituente di una categoria qualunque sulla base dello stesso principio, assegnandogli un'estensione (appropriata alla sua categoria) in ogni mondo possibile. L'idea è di Carnap; recentemente è stata applicata alla semantica delle lingue formalizzate con operatori intensionali, in scritti di diversi filosofi che si rifanno alla tradizione tarskiana e carnapiana.<sup>6</sup>

Per esempio, l'interpretazione di un nome dovrebbe darci in ogni mondo possibile la cosa nominata (se c'è). Essa può essere concepita come una funzione da mondi possibili a cose in essi contenute. Un'interpretazione per un predicato a un argomento dovrebbe dare le cose riguardo alle quali quel predicato è vero in un qualunque mondo possibile dato; può essere concepita

---

<sup>6</sup>Cfr. Rudolf Carnap, *Meaning and Necessity*, University of Chicago Press, Chicago, II edizione, 1956, pp. 181-182; Jakk Hintikka, "Modality as Referential Multiplicity", *Eripainos Ajatus*, 20 (1957), pp. 49-64; Saul Kripke, "Semantical Considerations on Modal Logic", *Acta Philosophica Fennica*, 16 (1963), pp. 83-94 [trad. it. di E. Bencivenga, "Considerazioni semantiche sulla logica modale", in *Riferimento e modalità*, cit. (N.d.T.)]; David Kaplan, *Foundations of Intensional Logic*, University Microfilms, Ann Arbor, 1964; Richard Montague, "Pragmatics", *Contemporary Philosophy - La Philosophie Contemporaine*, a cura di Raymond Klibansky, La Nuova Italia, Firenze, 1968; Montague, "On the Nature of Certain Philosophical Entities", *The Monist*, 53 (1969); Dana Scott, "Advice on Modal Logic", presentato al Free Logic Colloquium svoltosi alla University of California a Irvine, maggio 1968 [ora in *Philosophical Problems in Logic*, a cura di K. Lambert, Reidel, Dordrecht, 1970. (N.d.T.)] Montague e Scott propongono un trattamento unificato dell'intensione e dell'indicalità in cui le estensioni vengano assegnate relativamente a punti di riferimento, cioè a combinazioni di un mondo possibile con più tratti pertinenti del contesto: tempo, luogo, parlante, uditorio, ecc. Montague dà al suo lavoro il nome di pragmatica a causa di tale relativizzazione al contesto; ma esso non ha a che fare con quella specie di considerazioni pragmatiche che determinano quale lingua possibile è usata da una data popolazione.

come una funzione da mondi a insiemi di cose in essi contenute (o eventualmente, se non c'è niente che si trovi in più di un mondo, come un unico insieme contenente cose provenienti da mondi diversi). Un'interpretazione per un predicato a  $n$  argomenti può essere concepita come una funzione da mondi a insiemi di  $n$ -uple di cose in essi contenute. È possibile fornire di interpretazione anche quei costituenti che resistono al trattamento in una semantica referenziale limitata al mondo attuale. L'interpretazione di un operatore modale, per esempio, potrebbe essere concepita come una funzione da mondi possibili a insiemi di condizioni di verità: cioè a insiemi di insiemi di mondi possibili.

È questa specie di interpretazione - un'assegnazione di estensione in ogni mondo possibile - che, secondo la mia proposta, la grammatica dovrebbe associare ai costituenti, al fine di costruire per l'enunciato interpretazioni della specie che vogliamo. (Per l'attribuzione dei modi andrebbe preso, al livello della grammatica, qualche tipo di provvedimento particolare.) Ovviamente tali interpretazioni - come le estensioni in un mondo singolo - sono suscettibili di essere assegnate in base alle caratteristiche delle occasioni di emissione; per questo motivo, il significato di un enunciato non è una qualunque interpretazione, ma piuttosto la funzione secondo la quale l'interpretazione di esso (o l'insieme di interpretazioni alternative di esso) dipende dalle occasioni della sua emissione

### 5.3 Semantica in una lingua possibile

Con questo abbiamo completato il nostro esame delle lingue possibili. Possiamo definire alcune proprietà semantiche degli enunciati in relazione alle lingue possibili in generale, come abbiamo fatto in relazione alle lingue di segnalazione verbale. Ma queste proprietà, se vogliamo ammettere l'indicività e l'ambiguità, saranno in generale relazioni a quattro posti tra un enunciato  $\sigma$  di  $\mathcal{L}$ , la lingua  $\mathcal{L}$ , un'occasione possibile  $o$  di emissione di  $\sigma$  e un'interpretazione  $\langle \mu, \tau \rangle$  assegnata da  $\mathcal{L}$ , a  $\sigma$  in  $o$ . Se  $\sigma$  è *eterno* in  $\mathcal{L}$  - se cioè  $\mathcal{L}$  gli assegna, in ogni occasione possibile della sua emissione, lo stesso insieme di interpretazioni - possiamo omettere la menzione dell'occasione  $o$ . Se  $\sigma$  è *non ambiguo* in  $\mathcal{L}$  in  $o$  - se cioè  $\mathcal{L}$  gli assegna, in  $o$ , un'unica interpretazione - possiamo omettere la menzione dell'interpretazione  $\langle \mu, \tau \rangle$ .

$\sigma$  è *indicativo* in  $\mathcal{L}$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$  e  $\mu$  è 0. (Analogamente per gli altri quattro modi nominati.)

$\sigma$  è *vero* in  $\mathcal{L}$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  vale nel - cioè contiene il - mondo possibile  $w$  in cui è situata l'occasione possibile  $o$  di emissione di  $\sigma$ .

$\sigma$  è *falso* in  $\mathcal{L}$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  non vale nel mondo possibile  $w$  in cui è situata l'occasione  $o$ .

Parlando della verità e della falsità degli enunciati eterni di  $\mathcal{L}$ , o faremo menzione di un mondo possibile  $w$  oppure parleremo del mondo attuale.

$\sigma$  è *vero* in  $\mathcal{L}$  nel mondo  $w$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se  $\mathcal{L}$  assegna a  $\sigma$ , in ogni occasione possibile della sua emissione, un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  vale nel mondo possibile  $w$ . (Analogamente per la falsità.)

$\sigma$  è *vero* in  $\mathcal{L}$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se  $\mathcal{L}$  assegna a  $\sigma$ , in ogni occasione possibile della sua emissione, un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  vale nel mondo attuale. (Analogamente per la falsità.)

Nel caso più semplice è possibile assegnare la verità in  $\mathcal{L}$  a un enunciato eterno non ambiguo di  $\mathcal{L}$ . In mancanza di stipulazioni contrarie, essa va concepita come verità nel mondo attuale.

$\sigma$  è *vero* in  $\mathcal{L}$  se e solo se  $\mathcal{L}$  assegna a  $\sigma$ , in ogni occasione possibile della sua emissione, un'unica interpretazione fissa  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  vale nel mondo attuale. (Analogamente per la falsità.)

Le condizioni di verità assegnate da  $\mathcal{L}$  possono essere universali, valendo in - cioè contenendo - ogni mondo possibile; oppure possono essere vuote, non valendo in alcun mondo possibile; oppure possono stare tra i due estremi, valendo in alcuni mondi ma non in altri. Secondo i casi, possiamo chiamare gli enunciati analitici, contraddittori o sintetici in  $\mathcal{L}$ , in occasioni e in interpretazioni.

$\sigma$  è *analitico* in  $\mathcal{L}$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  vale in ogni mondo possibile.

$\sigma$  è *contraddittorio* in  $\mathcal{L}$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo

se  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  non vale in alcun mondo possibile.  $\sigma$  è *sintetico* in  $\mathcal{L}$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  vale in alcuni mondi possibili ma non in altri.

Di nuovo, è possibile semplificare i definendo quando si parla di enunciati che sono eterni in  $\mathcal{L}$ , non ambigui in  $\mathcal{L}$  in  $o$ , o entrambe le cose.

$\sigma$  è *analitico* in  $\mathcal{L}$  se e solo se  $\mathcal{L}$  assegna a  $\sigma$ , in ogni occasione possibile della sua emissione, un'unica interpretazione fissa  $\langle \mu, \tau \rangle$  e la condizione di verità  $\tau$  vale in ogni mondo possibile. (Analogamente per la contraddittorietà e la sinteticità.)

Si osservi che un enunciato indicale non ambiguo può essere vero in  $\mathcal{L}$  in ogni occasione possibile della sua emissione senza essere analitico in  $\mathcal{L}$  in alcuna occasione. Gli enunciati che si ritiene che abbiano queste prerogative in italiano comprendono “Io adesso sono qui”, “Io esisto”, “Io sono sveglio”, “Io sto dicendo qualcosa”, e simili.<sup>7</sup> A un enunciato di questo tipo, in qualunque occasione possibile della sua emissione, viene assegnata una condizione di verità che vale nel mondo in cui l'occasione di emissione è situata, ma non vale in qualche altro mondo possibile. È rassicurante scoprire che non abbiamo confuso queste prerogative con l'analiticità, come sarebbe stato fin troppo facile fare.

C'è una possibilità opposta che non è rassicurante e che giunge inaspettata. Un enunciato indicale non ambiguo può essere analitico in un'occasione possibile della sua emissione e falso in un'altra. Si consideri l'enunciato italiano “È un quadrato perfetto”. Prendiamo un'occasione della sua emissione in cui l'unica entità in questione nel discorso precedente era il numero 49, al quale ci si riferiva mediante il numerale “49”. In questo caso penso che la condizione di verità assegnata al nostro enunciato in quell'occasione valga

---

<sup>7</sup>Tutti questi esempi familiari sono soggetti a obiezioni. 1) “Qui” “ora” e il tempo presente possono riferirsi, in alcuni casi, non al luogo e al momento dell'occasione di emissione bensì al luogo e al momento dell'occasione di ascolto voluta. 2) Un fantasma che affermasse di esistere potrebbe affermarlo falsamente, se c'è un'accezione in cui esistere - per un essere umano o ex-umano - significa essere vivo. 3) Può darsi che un uomo che parla nel sonno riesca a eseguire l'azione di emettere qualcosa. Cito queste obiezioni non per sostenerle o per confutarle, ma solo per dire che mi interessa lo *status* che è stato attribuito ai miei esempi - non il problema se sono davvero dei buoni esempi di quello status.

in ogni mondo possibile in cui 49 è un quadrato perfetto, cioè in ogni mondo possibile. Si tratta quindi della condizione di verità universale. L'enunciato è analitico in italiano in quell'occasione. Ma prendiamo un'altra occasione della sua emissione in cui l'entità in questione era il numero 48, al quale ci si riferiva mediante il numerale "48". In *quella* occasione il nostro enunciato non è analitico in italiano; è falso, anzi contraddittorio, in italiano.

In ciò non trovo niente di sbagliato. Ricordiamo che la maggior parte delle discussioni sull'analiticità ha ignorato gli enunciati analitici indicali; quindi non dovremmo avere alcuna aspettativa già consolidata riguardo a essi. Inoltre, sebbene il nostro enunciato sia analitico in italiano in certe occasioni, non è *simpliciter* analitico in italiano, in base alla definizione data; mentre alcuni altri enunciati indicali come "Ieri è passato" sono non solo analitici in italiano in occasioni della loro emissione, ma sono anche *simpliciter* analitici in italiano. Qualunque occasione possibile di emissione dell'enunciato "Ieri è passato" è situata in qualche giorno  $g$ . "La condizione di verità assegnata all'enunciato in quella occasione vale in ogni mondo possibile in cui il giorno  $g - 1$  precede il giorno  $g$ ; ma questa è la condizione di verità universale, indipendentemente dal  $g$  in questione. Al nostro enunciato viene quindi assegnata un'unica interpretazione fissa in ogni occasione possibile della sua emissione, e questa interpretazione contiene la condizione di verità universale.

Di solito può essere chiamato contraddittorio un enunciato di un modo qualunque; ma solo un enunciato indicativo può essere chiamato analitico. Io ho trascurato questa inutile restrizione: l'analiticità - come la contraddittorietà, la verità e la falsità - dipende dalla condizione di verità assegnata da  $\mathcal{L}$ , indipendentemente dal modo associato. Un imperativo analitico è: "Metti un cappello oppure non metterlo!" Un commissivo analitico è: "Prometto di rimanere scapolo finché non mi sposerò." Un permissivo analitico è: "Puoi mentire tutte le volte che dici il falso." (Per ragioni che dovremmo considerare più avanti, dubito che ci sia qualcosa come una domanda analitica.)

Con questo è terminata la mia descrizione veloce, ma molto generale, delle lingue possibili e delle proprietà semantiche dei loro enunciati.

Passiamo adesso al problema seguente: se  $\mathcal{L}$  è una lingua possibile e  $P$  è una popolazione di agenti, quale relazione deve sussistere tra  $\mathcal{L}$  e  $P$  perché si verifichi che  $\mathcal{L}$  sia una lingua attuale di  $P$ ? La risposta ovvia è: i membri di  $P$  devono usare la lingua  $\mathcal{L}$ . Ma questa risposta non è né informativa né chiara. Essi devono usare  $\mathcal{L}$  in un certo modo.

Se ciascuno in  $P$  usasse  $\mathcal{L}$  per dire bugie o per cantare pezzi d'opera (senza capire le parole), userebbe  $\mathcal{L}$  ma non nel modo giusto;  $\mathcal{L}$  non sarebbe

la sua lingua. Soltanto quando sappiamo già che cosa fa di una lingua la lingua attuale di una popolazione possiamo dire quale specie di uso avevamo in mente.

La risposta successiva è questa: i membri di  $P$  devono dare agli enunciati di  $\mathcal{L}$  le interpretazioni (in occasioni di emissione) che  $\mathcal{L}$  assegna a questi enunciati. Ma questa è un'altra risposta non informativa. Che specie di azione è il dare un'interpretazione a un enunciato? Non un'azione che chiunque possa fare semplicemente proponendosi di farla. Non posso dire "Qui c'è freddo" e intendere "Qui c'è caldo"<sup>8</sup> - perlomeno, non senza un po' d'aiuto da parte dei miei amici.

## 5.4 Convenzioni di veracità

Stiamo cercando di scoprire che cosa devono fare i membri di una popolazione  $P$  per far sì che una certa lingua possibile  $\mathcal{L}$  sia la loro lingua attuale: cioè per usare  $\mathcal{L}$ , nel senso che stiamo cercando di precisare. Sicuramente si tratta di qualcosa che essi fanno in conformità a una convenzione: qualcosa che ciascuno in  $P$  fa perché si aspetta che anche i suoi compagni di conversazione in  $P$  lo facciano, e perché un comune interesse a comunicare lo spinge a desiderare di svolgere la propria parte se essi svolgono la loro. Questo lo sappiamo appunto perché sappiamo che importa poco (a lungo andare) quale lingua usiamo, purché usiamo la stessa.

Sosterrò che la convenzione di linguaggio per mezzo della quale i membri di una popolazione  $P$  usano una data lingua possibile  $\mathcal{L}$  può essere descritta nel modo migliore come convenzione di veracità in  $\mathcal{L}$ .<sup>9</sup> Questa è la conclusione cui siamo già arrivati a proposito del caso particolare in cui  $\mathcal{L}$  è una lingua di segnalazione verbale. In quel caso abbiamo dedotto l'esistenza di una convenzione di veracità in  $\mathcal{L}$  (1) dalla nostra precedente spiegazione delle convenzioni di segnalazione e (2) dalla nostra applicazione del concetto

---

<sup>8</sup>Ludwig Wittgenstein, *Philosophical Investigations*, Blackwell, Oxford, 1958, sez. 510. [Trad. it. di R. Piovesan e M. Trinchero, *Ricerche filosofiche*, Einaudi, Torino, 1967. (N.d.T.)]

<sup>9</sup>Sono debitore di questa proposta a Erik Stenius, "Mood and Language Game", *Synthese*, 17 (1967), pp. 254-274. Egli propone che l'uso del linguaggio sia codificato da regole che danno le condizioni di verità dei radicali di enunciato (enunciati meno gli indicatori di modo) e da regole che prescrivono la specie opportuna di veracità per ciascun modo. Egli prende in considerazione tre modi: indicativi, imperativi e interrogativi del tipo sì-no. Io ho adattato la sua proposta inserendo le condizioni di verità nell'identificazione delle lingue possibili e interpretando le sue regole di veracità come convenzioni di veracità. Stenius, a sua volta, si dichiara debitore al dott. H. Johansen di Copenaghen dell' "idea che dire la verità possa essere considerato, in un certo senso, come una regola semantica."

di verità alla segnalazione convenzionale. Adesso esamineremo in che modo anche lingue meno ristrette possono venir usate in conformità a convenzioni di verità.

Rimandiamo a più tardi l'esame dei modi non indicativi, dell'indicalità e dell'ambiguità. Abbiamo dunque una lingua  $\mathcal{L}$  che assegna a ciascuno dei suoi enunciati un'unica interpretazione al modo indicativo, la stessa in ogni occasione possibile di emissione di quell'enunciato. Essere verace in  $\mathcal{L}$  significa cercare di non emettere alcun enunciato di  $\mathcal{L}$  che non sia vero in  $\mathcal{L}$ ; ma non significa soltanto questo. Non si può essere veraci in  $\mathcal{L}$  semplicemente non emettendo mai enunciati di  $\mathcal{L}$ . (In questo momento io non sono verace in gallese.) Essere verace in  $\mathcal{L}$ , in un'accezione più positiva, significa partecipare agli scambi verbali con l'occasionale emissione di enunciati di  $\mathcal{L}$ , e insieme cercare di non emettere nulla che non sia vero in  $\mathcal{L}$ . Si può naturalmente rimanere in silenzio in qualunque particolare momento di una conversazione senza smettere di essere veraci in  $\mathcal{L}$ ; ma state zitti troppo a lungo - o a un prezzo troppo alto per i vostri fini sociali - e vi trasformerete gradualmente, da utente verace di  $\mathcal{L}$  (che non ha voglia di parlare), in un non utente di  $\mathcal{L}$ .

Rispetto alle lingue di segnalazione verbale non abbiamo avuto bisogno di distinguere tra verità positiva e negativa, in quanto la convenzione di segnalazione non permetteva di scegliere tra il parlare e lo stare zitti. Per conformarsi bisognava emettere qualunque enunciato della lingua fosse vero. Ma sarebbe assurdo richiedere che le nostre convenzioni di linguaggio proibiscano il silenzio in un qualunque momento di una conversazione ordinaria. Eppure non si può dire che chi sceglie *sempre* il silenzio si conforma alle nostre convenzioni di linguaggio.

Se la nostra semplice lingua  $\mathcal{L}$  è la lingua attuale di una popolazione  $P$ , scopriremo certamente una regolarità  $R$  di verità in  $\mathcal{L}$ . I membri di  $P$  si scambieranno emissioni di enunciati di  $\mathcal{L}$ , e cercheranno quasi sempre di evitare di emettere enunciati non veri in  $\mathcal{L}$ . (A volte mentiranno; ma possiamo tollerare delle eccezioni a una regolarità convenzionale. A volte si sbaglieranno; ma la regolarità cui penso consiste nel *cercare* di essere verace in  $\mathcal{L}$ , non nel riuscirci. Le convenzioni sono regolarità nella scelta di un'azione.) Inoltre, sarà di conoscenza comune, in  $P$ , il fatto che i membri di  $P$  cercano quasi sempre di conformarsi a  $R$ . Che dire delle altre condizioni di convenzionalità di  $R$ ?

Qualunque membro di  $P$  vorrà (quasi sempre) conformarsi a  $R$  se lo fanno gli altri membri; anche questi ultimi vorranno che egli lo faccia; e queste preferenze saranno conoscenza comune in  $P$ . Possiamo senz'altro assumere che  $\mathcal{L}$  sia una lingua utile, perché altrimenti non potrebbe essere la lingua

attuale di *P*. La veracità uniforme in  $\mathcal{L}$  rende allora possibile il buon esito della comunicazione, che risponde a tutti i generi di interesse dei membri di *P*. Consideriamo una qualunque conversazione ordinaria tra membri di *P*. Qualcuno vuole far sì che gli altri condividano certe sue credenze. Non importa il motivo per cui egli lo desidera: può darsi che voglia che essi agiscano in base a quelle credenze, per ragioni egoistiche o altruistiche; può darsi che cerchi di educarli; può darsi che stia soltanto facendo passare il tempo. Egli può conseguire il suo scopo emettendo enunciati che ritiene veri in  $\mathcal{L}$ , cioè conformandosi opportunamente a *R*. E il motivo per cui può conseguire il suo scopo conformandosi a *R* è che altri si sono conformati a *R* nel passato: ossia coloro che hanno dato forma alle abitudini e alle attese del suo attuale uditorio. A causa della prevalente regolarità di veracità in  $\mathcal{L}$ , il suo uditorio si è abituato alla veracità in  $\mathcal{L}$ . Essi sono abitualmente veraci in  $\mathcal{L}$ . Abitualmente agiscono ed elaborano credenze in base all'assunzione che gli altri siano veraci in  $\mathcal{L}$ . Essendo pratici di  $\mathcal{L}$ , lo fanno velocemente con facilità. In questo modo i membri di *P* hanno normalmente motivo di essere veraci in  $\mathcal{L}$ , subordinatamente alla recente diffusa veracità in  $\mathcal{L}$  da parte dei membri di *P*.

In base a questa spiegazione, la coordinazione raggiunta per mezzo della veracità in  $\mathcal{L}$  è una coordinazione, a una parte, diffusa tra comunicanti che, in tempi diversi, comunicano con lo stesso uditorio. Ogni comunicante desidera essere verace in  $\mathcal{L}$  perché è questo che i comunicanti precedenti hanno indotto l'uditorio ad aspettarsi. Un membro dell'uditorio, in quanto tale, non è vincolato dalla convenzione. Egli semplicemente ascolta ed eventualmente elabora delle credenze, nella consapevolezza che probabilmente il comunicante è verace in  $\mathcal{L}$ . Solo quando toccherà a lui fare la parte del comunicante si conformerà alla convenzione di veracità in  $\mathcal{L}$ .

Abbiamo descritto diversamente il caso di una lingua di segnalazione verbale con segnali indicativi. La convenzione di veracità era a due parti: il comunicante cercava di emettere un enunciato vero nella lingua e l'uditorio agiva nel modo che sembrava migliore nell'ipotesi che egli fosse riuscito nel suo intento. La coordinazione di cui si andava alla ricerca era la coordinazione tra comunicante e uditorio.

Penso che sia giusto trattare i due casi diversamente. La segnalazione verbale viene effettuata avendo in mente un fine ben determinato, ed è conoscenza comune di quale fine si tratta. Ma questo non è il caso generale. Io non vi dico soltanto ciò che avete bisogno di sapere proprio ora al fine di contribuire a realizzare i nostri scopi comuni. Spesso non c'è niente di particolare che l'uditorio debba fare se il comunicante ha detto la verità. L'uditorio deve eventualmente elaborare una credenza, ma nor-

malmente questa non è un'azione volontaria e quindi non è conforme a una convenzione. Se anche l'uditorio dovesse agire, l'azione potrebbe non essere rispondente a un interesse comune al comunicante e all'uditorio, potrebbe non essere l'azione che il comunicante voleva determinare, e così via; e quindi potrebbe non essere descrivibile in modo plausibile come azione regolata da una convenzione. Indubbiamente esiste uno spettro continuo che va dalla segnalazione verbale alla chiacchierata futile, e coordinazioni a due parti e a una parte possono essere mescolate in varie proporzioni. Ma la generalità si raggiunge concentrandosi sulla coordinazione a una parte tra comunicanti, la quale è presente in ogni comunicazione indicativa convenzionale, e non sull'occasionale coordinazione a due parti tra comunicante e uditorio.

Infine, una convenzione deve ammettere un'alternativa. Se  $R$  è una convenzione di veracità in  $\mathcal{L}$ , le sue alternative sono regolarità di veracità in opportune lingue possibili che siano alternative. Per diversi motivi, non ogni altra lingua possibile è un'alternativa opportuna. Ricordiamo che molte entità bizzarre sono ancora chiamate lingue possibili perché non ci siamo sobbarcati alla fatica di escluderle. Non saremo capaci neppure di definire la veracità rispetto a una lingua possibile arbitraria, ma solo rispetto a una lingua che si limita ai primi pochi modi e che non presenta un grado eccessivo di ambiguità. Anche in questo caso, avremo lingue prive di grammatica, lingue banali, lingue impronunciabili, lingue non acquisibili, lingue i cui enunciati più brevi richiedono ore per essere emessi, lingue in cui nessun enunciato dice nulla di ciò che si vorrebbe dire ecc. Una lingua alternativa a  $\mathcal{L}$  deve essere una lingua possibile rispetto alla quale la veracità è definibile, e una lingua abbastanza conveniente e utile da essere seriamente candidata all'uso da parte dell'uomo. Ma non deve essere troppo vicina a  $\mathcal{L}$ , in quanto deve essere quasi sempre impossibile agire in conformità sia a una convenzione che alla sua alternativa. Se  $\mathcal{L}$  e le sue alternative condividessero alcuni enunciati, e assegnassero a quegli enunciati condizioni di veracità in parte coincidenti, potrebbe diventare possibile essere veraci nello stesso tempo in  $\mathcal{L}$  e nelle sue alternative. Di nuovo, non dobbiamo temere che la convenzione di veracità in  $\mathcal{L}$  manchi di alternative. Ci sono tutte le lingue attuali di popolazioni diverse, e numerosissime altre lingue possibili che andrebbero altrettanto bene di quelle attuali.

Abbiamo così completato il nostro esame delle condizioni definitorie che devono essere soddisfatte perché in  $P$  ci sia la convenzione di essere veraci nella lingua semplice  $\mathcal{L}$ ; e abbiamo scoperto che, nell'ipotesi che  $\mathcal{L}$  sia la lingua attuale di  $P$ , queste condizioni sarebbero soddisfatte. In  $P$  prevarrebbe una convenzione di veracità in  $\mathcal{L}$ . Questa convenzione sarebbe sostenuta da un certo tipo di interesse da parte dei membri di  $P$ : un interesse alla comu-

nicazione, al fatto di essere in grado di controllare reciprocamente, in una certa misura, le proprie credenze e le proprie azioni per mezzo di suoni e di segni.

Alcuni sono portati a negare che ci possa essere una convenzione di veracità in  $\mathcal{L}$ , argomentando che la veracità è un obbligo morale. Questa obiezione si fonda in parte sulla confusione tra veracità in una data lingua possibile  $\mathcal{L}$  e veracità nella lingua di una popolazione, qualunque lingua possibile questa possa essere. È la prima, non la seconda, che è convenzionale. Entrambe sono obbligatorie, secondo le nostre comuni opinioni morali. Io sono obbligato a essere verace in  $\mathcal{L}$ , a condizione che i miei compagni siano veraci in  $\mathcal{L}$ ; sono obbligato a essere verace in qualunque lingua usino, a condizione che usino effettivamente l'una o l'altra lingua. Ammetto che questi sono obblighi morali, ma nego che impediscano alla veracità in  $\mathcal{L}$  di essere convenzionale.

Perché potrebbe sembrare il contrario? Suppongo che sia perché le convenzioni, quali sono state definite, poggiano sulla preferenza. Conformarsi a una convenzione di veracità in  $\mathcal{L}$  significa essere verace in  $\mathcal{L}$  perché ciò corrisponde alle proprie preferenze per tale comportamento, visto che gli altri si comportano così. Ma le persone virtuose si sentono vincolate dall'obbligo di essere veraci in  $\mathcal{L}$  se lo sono gli altri, anche se ciò è contrario alle loro preferenze.

È vero questo: non è possibile che uno che è verace in  $\mathcal{L}$  contro le proprie preferenze stia agendo in conformità a una convenzione. Ma casi simili sono eccezionali. Nel mondo quale noi lo conosciamo - e quale deve essere, se il linguaggio deve continuare a essere usato da uomini macchiati dal peccato - quasi ognuno ha quasi sempre motivo di far sì che gli altri condividano le proprie credenze, e quindi ha motivo di conformarsi alle convenzioni di veracità. Così, nel caso normale, può darsi che si adempiano degli obblighi morali e nello stesso tempo si agisca in base alle proprie preferenze. Ogniqualvolta ciò avviene, possiamo avere un caso di conformità a una convenzione. Soltanto alcuni casi eccezionali richiedono di essere descritti diversamente.

Una convenzione di veracità in  $\mathcal{L}$  è tanto un contratto sociale quanto una convenzione. Non solo ciascuno preferisce la veracità in  $\mathcal{L}$  da parte di tutti alla veracità in  $\mathcal{L}$  da parte di tutti tranne lui. C'è di più: ciascuno preferisce la veracità uniforme in  $\mathcal{L}$  alla Babele, allo stato di natura. Così ciascuno dovrebbe riconoscere l'obbligo di correttezza consistente nel ricambiare i benefici che gli sono derivati dall'altrui veracità in  $\mathcal{L}$ , essendo egli stesso verace in  $\mathcal{L}$ . Nei casi eccezionali questo obbligo sarà il suo unico motivo di essere verace in  $\mathcal{L}$ . Nei casi normali esso sarà presente ma superfluo, in quanto egli avrà anche sufficienti motivi di interesse personale.

Un'obiezione diversa può essere mossa alla mia opinione secondo cui le convenzioni di linguaggio sono convenzioni di veracità in una data lingua possibile  $\mathcal{L}$ . Ricordiamo che cos'è  $\mathcal{L}$ : una certa funzione i cui argomenti sono coppie costituite da un enunciato (una sequenza finita di tipi di suoni o di caratteri) e da un'occasione possibile della sua emissione (una coppia costituita da un mondo possibile e da una posizione spazio-temporale in esso) e i cui valori sono insiemi di interpretazioni (coppie costituite da un numero di codice e da un insieme di mondi possibili). Ora, è incredibile che qualunque utente ordinario di  $\mathcal{L}$  abbia nozione di un'entità così complicata. Come può dunque partecipare a una convenzione relativa a  $\mathcal{L}$ ? Come può avere attese e preferenze relative alla veracità in  $\mathcal{L}$ ?

La mia risposta è che può averle *in sensu diviso*; per far ciò egli non ha bisogno di condividere il nostro complicato concetto di lingua possibile. Non ha neppure bisogno di condividere i nostri concetti di funzione, sequenza, insieme, coppia e mondo possibile. Può essere un filosofo nominalista che non crede in nessuna di queste cose oppure un bifolco che non ne ha mai sentito parlare. Tutto quello che deve fare è arrivare ad avere le aspettative e le preferenze giuste riguardo a casi particolari di quella che noi - ma non lui - chiameremo veracità in  $\mathcal{L}$ . E questi casi particolari sono pensieri e azioni umane assolutamente banali. Gli unici a usare esplicitamente concetti esoterici siamo noi, quando compendiamo una massa enorme e eterogenea di azioni determinate dalle nostre credenze classificandole in base a una descrizione inventata: veracità in  $\mathcal{L}$ .

È dunque logicamente possibile conformarsi a una convenzione di veracità in  $\mathcal{L}$  senza avere un concetto generale di veracità in  $\mathcal{L}$ . Questo non equivale a dire che è *psicologicamente* possibile farlo senza avere qualcosa di molto analogo a quel concetto; né io voglio dire una cosa del genere. L'utente di  $\mathcal{L}$  è un essere finito con un'esperienza molto limitata; e tuttavia ha acquisito, in qualche modo, un repertorio enorme, ed enormemente vario, di inclinazioni all'azione, all'aspettativa e alla preferenza in una grande varietà di situazioni. Ecco che cosa gli occorre per essere - come diremmo noi, ma non lui - solitamente verace in  $\mathcal{L}$  e abituato ad aspettarsi la veracità in  $\mathcal{L}$  da parte degli altri. Di questa competenza non riesco a immaginare alcuna spiegazione che non postuli una qualche specie di sostituto inconscio della nozione generale di veracità in  $\mathcal{L}$ .

Consideriamo un ciclista. Qualunque teoria credibile della sua competenza gli attribuirà dei processi mentali inconsci strettamente analoghi all'uso dei concetti generali della fisica, quando si sia a conoscenza delle leggi generali della fisica. Ma, a meno che il nostro ciclista sia anche un fisico, sarebbe sbagliato affermare che egli possedeva quei concetti o conosceva quelle leggi,

anche se le sue attese relative a una grande varietà di situazioni particolari sono in armonia con quelle leggi.

Ho così completato la spiegazione delle convenzioni di veracità in lingue semplici prive di modi non indicativi, di indicività e di ambiguità. Rimandando ancora l'indicività e l'ambiguità, passiamo alla veracità in lingue con imperativi, interrogativi, commissivi e permissivi.

Ricordando la discussione sugli imperativi nelle lingue di segnalazione verbale, possiamo concepire, grosso modo, la veracità imperativa come obbedienza. Cioè, spetta all'uditorio rendere veri in  $\mathcal{L}$  gli enunciati imperativi di  $\mathcal{L}$ , cercando di comportarsi in modo che le condizioni di verità loro assegnate da  $\mathcal{L}$  valgano.

Come prima approssimazione, potremmo considerare questa formulazione di una convenzione di veracità in  $\mathcal{L}$  tra membri di  $P$ , nella sua applicazione agli imperativi: se un qualunque membro di  $P$  pensa di essere stato l'uditorio cui era destinata l'emissione, da parte di un altro membro di  $P$ , di un enunciato imperativo  $\sigma$  di  $\mathcal{L}$ , allora egli cerca di rendere  $\sigma$  vero in  $\mathcal{L}$ .

Non è un'approssimazione molto buona. Certamente non esiste tra noi la convenzione di rendere vero qualunque imperativo. Ci sono i cattivi consigli, gli ordini non autorizzati, le pretese esorbitanti, le richieste improponibili ecc. Sono tutti imperativi che possiamo ignorare senza violare una convenzione. È solo quando il comunicante sta in una certa *relazione di autorità* con l'uditorio cui si rivolge che quest'ultimo violerà la convenzione se non cercherà di rendere veri gli imperativi del comunicante. Almeno in molti casi, e forse in tutti, l'opportuna relazione di autorità è la seguente: il comunicante e il suo uditorio hanno un comune interesse a rendere possibile al comunicante il controllo delle azioni dell'uditorio almeno entro una certa sfera, per mezzo dell'emissione di espressioni verbali, e questo fatto costituisce, tra loro, una conoscenza comune.

Per la nostra definizione di problema di segnalazione verbale, questa relazione vale tra il comunicante e l'uditorio in qualunque problema di segnalazione. Essa vale ogniqualvolta il comunicante dà consigli all'uditorio ed è conoscenza comune che egli è in condizione di dare dei buoni consigli e che intende farlo. Vale ogniqualvolta il comunicante dà all'uditorio un ordine nel corso dell'esecuzione di qualche lavoro d'equipe, se è conoscenza comune che è questo che egli sta facendo. Vale ogniqualvolta il comunicante fa una richiesta non esorbitante ed è conoscenza comune che l'uditorio desidera fargli piacere.

Ma questa relazione di autorità non sembra valere tra un comunicante e un uditorio quando l'imperativo del comunicante è una richiesta fatta valere con una minaccia. Vale l'opposto: è conoscenza comune che l'uditorio

sarebbe in una condizione più agevole se il comunicante non fosse in grado di controllare le sue azioni. Forse un uditorio che cede a una richiesta non dovrebbe essere descritto come un esempio di conformità a una convenzione di veracità, per quanto possiamo ammettere che la richiesta per mezzo di imperativi è possibile in virtù dell'esistenza di convenzioni di veracità che regolano quegli imperativi in altre situazioni, situazioni in cui il comunicante si trova nella relazione d'autorità con l'uditorio. O forse dovremmo dire che l'analisi migliore delle richieste consiste nell'interpretarle non come imperativi ma come commissivi - minacce - anche quando grammaticalmente sono simili a imperativi.

Siamo arrivati alla seguente formulazione di una convenzione di veracità in  $\mathcal{L}$ , nella sua applicazione agli imperativi. Se un membro  $x$  di  $P$  pensa di essere stato l'uditorio cui era destinata l'emissione, da parte di un altro membro  $y$  di  $P$ , di un enunciato imperativo  $\sigma$  di  $\mathcal{L}$ ; se  $y$  e  $x$  hanno un interesse comune a rendere possibile a  $y$  il controllo delle azioni di  $x$ , entro una certa sfera, per mezzo dell'emissione di espressioni verbali (e se ciò costituisce, tra loro, una conoscenza comune); se in quella sfera c'è qualche azione per mezzo della quale  $x$  può cercare di rendere  $\sigma$  vero in  $\mathcal{L}$ ; allora  $x$  cerca di rendere  $\sigma$  vero in  $\mathcal{L}$ .

Ho detto che una convenzione di veracità in  $\mathcal{L}$  impegnava il comunicante di un indicativo, ma non l'uditorio; analogamente, adesso sto descrivendo la convenzione come se impegnasse l'uditorio di un imperativo ma non il comunicante. Ma in questo caso non vedo alcun buon motivo per non includervi anche il comunicante: la sua parte consiste nell'emettere gli imperativi nel modo che gli sembra migliore, nell'ipotesi che il suo uditorio sia verace in  $\mathcal{L}$ . Cioè, quando si trova nella relazione di autorità con il suo uditorio, egli emette quegli imperativi che vuole che l'uditorio renda veri in  $\mathcal{L}$  e che l'uditorio può rendere veri in  $\mathcal{L}$  per mezzo di azioni che rientrano nella sfera accettabile di controllo. D'altra parte non sembra chiaramente necessario inserire la parte del comunicante nella formulazione della convenzione. Possiamo rappresentare altrettanto bene il comunicante come un agente che fa quello che gli sembra meglio, nella consapevolezza che la convenzione esiste e che probabilmente il suo uditorio si conformerà a essa. Come nel caso puramente indicativo, non è necessario che concentriamo la nostra attenzione sulla coordinazione a due parti e a breve scadenza tra un comunicante e il suo uditorio. C'è anche una coordinazione diffusa e a lungo termine tra tutti quelli che conversano con una data persona. Infatti, chiunque conversi con questa persona influisce sulle sue abitudini linguistiche e insieme consegue vari scopi conversazionali - o non riesce a conseguirli - in virtù delle abitudini linguistiche che ha già.

Passando agli interrogativi, possiamo risparmiarci qualche fatica. Invece di trattarli come un modo diverso, possiamo interpretarli come una specie di imperativi. (Perciò il numero di codice 2, riservato all'inizio di questo capitolo al modo interrogativo, non comparirà nelle interpretazioni assegnate dalle lingue che stiamo considerando.) Prendiamo la domanda "È pronto il pranzo?"; confrontiamola con l'imperativo "Dimmi se è pronto il pranzo!" o, più semplicemente, con "Dimmi 'sì' se il pranzo è pronto, 'no' altrimenti!" Oppure prendiamo la domanda "Qual è il tuo nome?"; confrontiamola con l'imperativo "Di' il tuo nome!" O infine prendiamo la domanda "Dimmi, chi scoprì la Groenlandia?"; confrontiamola con l'imperativo "Dimmi chi scoprì la Groenlandia!" Non riesco a trovare alcuna differenza rilevante in nessuna di queste coppie. Propongo quindi che le domande siano imperativi imperativi con argomento caratteristico, contrassegnati da una forma grammaticale caratteristica. Come qualunque imperativo, una domanda è resa vera in  $\mathcal{L}$  dall'esecuzione, da parte dell'uditorio, dell'azione ordinata: l'azione che chiamiamo dare una risposta vera alla domanda. Una domanda vera è una domanda la cui risposta è vera.<sup>10</sup> La veracità in  $\mathcal{L}$ , rispetto alle domande, consiste nel cercare di dare risposte vere, almeno quando colui che pone la domanda sta con qualcuno in una relazione di autorità. Questo è semplicemente un caso particolare di veracità in  $\mathcal{L}$  rispetto agli imperativi. Nello stesso tempo, è un caso particolare di veracità in  $\mathcal{L}$  rispetto agli indicativi; le risposte - almeno quando sono enunciati completi - sono indicativi, come pure azioni che rendono veri gli imperativi.

Ho detto, di sfuggita, che dubitavo dell'esistenza di domande analitiche. Se le domande sono imperativi, le domande analitiche sono imperativi che non possono non essere resi veri; infatti, gli enunciati di qualunque modo sono analitici in  $\mathcal{L}$  se e solo se viene loro assegnata la condizione di verità universale. Sono quindi domande alle quali non si può non dare una risposta vera. Ma qualunque domanda che si possa autenticamente dire tale ammette risposte alternative contrarie. Inoltre, si può non dare una risposta vera rimanendo in silenzio ed evitando così addirittura di rispondere.

Il modo commissivo delle promesse e delle minacce è affine all'indicativo. Spetta all'emittente di essere verace in  $\mathcal{L}$ , cercando di badare a che la

---

<sup>10</sup>Su questo punto non sono d'accordo con Stenius. In "Mood and Language Game" egli propone di considerare vero il radicale di enunciato di una domanda del tipo sì-no se e solo se "sì" è la risposta corretta. Egli prescrive quindi questa specie di veracità per le domande: "Rispondere alla domanda con 'sì' o 'no', secondo che il suo radicale di enunciato sia vero o falso" (p. 273). Questa è senza dubbio la nozione più naturale di verità per le domande del tipo sì-no; ma, diversamente dalla mia, non può essere estesa ad altre specie di domande.

condizione di verità del suo commissivo valga. Render vera in  $\mathcal{L}$  una promessa o una minaccia significa mantenerla, fare quello che si era detto che si stava per fare. Il mio indicativo “Tornerò” e il mio commissivo “Prometto di tornare” sono entrambi veri in italiano se e solo se io in seguito ritorno; e per essere verace in italiano io devo cercare di assicurare che emetto l’uno o l’altro enunciato solo se, in seguito, farò ritorno. In altre parole, devo cercare di agire in modo tale che ci sia una corrispondenza tra le mie parole di adesso e le mie azioni future. Ma ci sono due momenti diversi in cui posso cercare di assicurarmi questa corrispondenza: adesso, quando scelgo le mie parole, o in futuro, quando sceglierò le mie azioni.

Nel caso dell’indicativo “Tornerò”, essere verace significa cercare, *prima* di emettere l’enunciato, di fare in modo che le mie parole corrispondano alle mie azioni. Faccio questo cercando di prevedere o di decidere se tornerò, e astenendomi dal dire “Tornerò” se prevedo o decido di non tornare. Ai fini della mia veracità, non è pertinente che io continui o meno, dopo l’emissione, a cercare di garantire la corrispondenza. (È pertinente ai fini della verità del mio enunciato; ma essere verace significa cercare di rendere vero l’enunciato, e non necessariamente riuscirvi.) L’autopredizione avventata non è verace; la violazione imprevista di un’autopredizione no.

Nel caso del commissivo “Prometto che tornerò”, essere verace significa cercare, *dopo* aver emesso l’enunciato, di fare in modo che le mie parole corrispondano alle mie azioni. Faccio questo ricordando le mie parole e, di conseguenza, cercando di tornare. Ai fini della mia veracità, non è pertinente - sebbene sia pertinente ai fini della verità del mio enunciato - che io abbia o non abbia cercato, prima dell’emissione, di fare in modo che le mie parole corrispondessero alle mie azioni. Il mancato tentativo, imprevisto, di mantenere una promessa non è verace; una promessa avventata no.

Se i membri di una popolazione  $P$  sono veraci in  $\mathcal{L}$  e  $\sigma$  è un enunciato indicativo di  $\mathcal{L}$ , essi cercheranno, prima di emettere  $\sigma$ , di valutare se la condizione di verità assegnata da  $\mathcal{L}$  a  $\sigma$  vale. Se  $\sigma$  è un enunciato commissivo di  $\mathcal{L}$ , essi cercheranno, dopo aver emesso  $\sigma$ , di agire in maniera tale che valga la condizione di verità assegnata a  $\sigma$ . Supponiamo che tra loro valga una convenzione di veracità in  $\mathcal{L}$ . Allora, nel fare l’una o l’altra di queste cose, essi si conformano alla convenzione. Essi *possono* fare anche altre cose: possono cercare di rendere veri degli indicativi che sono già stati emessi, o cercare di non emettere commissivi che prevedono che non saranno resi veri. Ma queste azioni non sono regolate dalla convenzione di veracità in  $\mathcal{L}$ . Tra noi, esse non sono regolate proprio da nessuna convenzione.

È probabile che il lettore avverta di nuovo un disagio di cui abbiamo già discusso. La caratteristica essenziale della promessa e della minaccia,

come fanno gli strateghi consiste nell'impegnarsi a fare qualche cosa, che questa risulti, al momento, rispondente o meno alle nostre preferenze (in modo che le aspettative altrui relative alla nostra azione siano più stabili di quelle relative alle nostre preferenze). Ma un'azione normalmente conforme a una convenzione è, per definizione, un'azione rispondente alle preferenze di qualcuno. Così, come può una promessa essere mantenuta conformemente a una convenzione? Non può esserlo, quando si oppone veramente a una preferenza. Ma quando la preferenza concorda con l'obbligo implicito nella promessa, non c'è alcun motivo per cui il mantenimento di una promessa non debba essere conforme a una convenzione. Quest'ultimo è il caso normale, e deve valere in qualunque popolazione con una convenzione di veracità in una lingua contenente commissivi. Ciò non dovrebbe sorprenderci: noi normalmente manteniamo le promesse perché non vogliamo deludere le altrui legittime aspettative; perché ne trarremo vantaggio, visto quello che gli altri hanno fatto nell'aspettativa che la promessa venga mantenuta; perché non vogliamo incorrere in rappresaglie, distruggere la nostra reputazione per quanto riguarda la fedeltà alla parola data, o indebolire in generale la fiducia nelle promesse. Normalmente preferiamo mantenere le promesse. E questa è una preferenza condizionale, della specie richiesta affinché il mantenimento delle promesse sia un'azione conforme a una convenzione. Se gli altri non cercassero generalmente di rendere veri in  $\mathcal{L}$  gli enunciati commissivi di  $\mathcal{L}$  che hanno emesso, nessuno avrebbe alcuno dei motivi che ho menzionato per preferire di cercare di rendere veri in  $\mathcal{L}$  gli enunciati commissivi di  $\mathcal{L}$  che ha emesso. Qualunque promessa ha il potere di impegnare un uomo d'onore contro le sue preferenze. Ma la maggior parte delle promesse non ha mai richiesto l'esercizio di quel potere.

Gli enunciati permissivi hanno l'effetto di neutralizzare gli imperativi e i commissivi. Se permetto a qualcuno di fare qualcosa (o di lasciare qualcosa di non fatto), di solito cancello qualche comando che gli ho rivolto o qualche promessa che egli mi ha fatto. L'altro può approfittare del mio permesso per comportarsi in un modo che altrimenti sarebbe stato non verace, quindi contrario alla nostra convenzione quale essa si applica al mio comando o alla sua promessa iniziali. A causa del mio permesso, il suo comportarsi in quel modo non è non verace. La cancellazione può essere completa o parziale: dopo averti ordinato (o dopo aver ricevuto da te la promessa) di non calpestare il mio prato, posso dirti che d'ora in poi puoi calpestarlo tutte le volte che vuoi, oppure posso dirti che puoi calpestarlo soltanto per questa volta.

Non descriverò una nuova specie di veracità in  $\mathcal{L}$ , la veracità rispetto ai permissivi. I permissivi entreranno piuttosto nella formulazione di una con-

venzione di veracità in 2 per mezzo di una elaborazione della mia spiegazione della veracità rispetto agli imperativi e ai commissivi. Ciò non significa che il modo permissivo stesso verrà eliminato. Gli enunciati permissivi di  $\mathcal{L}$ , con le loro interpretazioni contenenti il numero di codice 4, rimarranno. Ma la nostra formulazione di una convenzione di veracità in  $\mathcal{L}$  non conterrà alcuna parte dedicata esclusivamente ai permissivi. Anzi, non è necessario aggiungere nulla alla nostra formulazione della convenzione. Dobbiamo soltanto fare un'osservazione relativa alla natura delle condizioni di verità che possono essere assegnate agli imperativi e ai commissivi in una lingua contenente i permissivi.

La condizione di verità assegnata da  $\mathcal{L}$  a un enunciato permissivo  $\sigma$  ci dà lo stato di cose che è permesso che valga quando  $\sigma$  è emesso: lo stato di cose in cui, per così dire, qualcuno approfitta del permesso che gli è stato concesso per mezzo di un'emissione di  $\sigma$ . Sappiamo così che cosa vuoi dire che un permissivo che è stato emesso sia reso vero in  $\mathcal{L}$ .

La specie più semplice di condizione di verità per un imperativo o un commissivo è l'insieme contenente esattamente quei mondi possibili in cui una certa persona esegue una certa azione. Possiamo chiamarla una condizione di verità *positiva*. Ma se un imperativo o un commissivo è condizionale "Prendi l'ombrello se c'è brutto tempo!" o "Ti prometto di aiutarti se avrai bisogno di me" - c'è anche una condizione di verità *negativa* contenente esattamente quei mondi in cui lo stato di cose antecedente non si verifica. Più sono gli antecedenti, più sono le condizioni di verità negative diverse. La condizione di verità complessiva di un imperativo o di un commissivo è l'unione di tutte le sue condizioni di verità positive e negative. Nel caso di un imperativo o di un commissivo che possa essere neutralizzato da un permesso, abbiamo un'altra specie di condizione di verità negativa contenente esattamente quei mondi in cui la neutralizzazione ha avuto luogo. Il mio progetto consiste quindi nel concepire qualunque imperativo o commissivo che possa essere neutralizzato come se fosse un condizionale esplicito, cioè come se fosse esplicitamente subordinato - *inter alia* - all'assenza di qualunque permesso neutralizzante.

Più precisamente, supponiamo di avere un imperativo o un commissivo la cui condizione di verità positiva riguarda l'azione di una certa persona  $x$ . Questo imperativo o commissivo può avere anche una condizione di verità negativa che può essere descritta nel modo seguente: l'insieme di tutti i mondi possibili in cui un qualunque enunciato permissivo  $\sigma$  di  $\mathcal{L}$  viene indirizzato a  $x$  da parte di chiunque si trovi in una certa posizione, e in cui  $\sigma$  viene successivamente reso vero in  $\mathcal{L}$  da  $x$ . (La neutralizzazione ha luogo se la condizione di verità di  $\sigma$  non è completamente inclusa nella condizio-

ne di verità positiva dell'imperativo o del commissivo originale; altrimenti otteniamo un mondo che appartiene sia alla nostra condizione di verità negativa che alla condizione di verità positiva iniziale.) Questa condizione di verità negativa sarà inclusa nell'unione delle condizioni di verità positive e negative che costituisce la condizione di verità complessiva dell'imperativo o del commissivo originale.

Non è necessario prendere provvedimenti particolari riguardo alle condizioni di verità negative di questa specie, ma soltanto osservare che sono possibili. E se qualche imperativo o commissivo di  $\mathcal{L}$  - non necessariamente tutti - ha tali condizioni di verità, l'uso degli opportuni permissivi di  $\mathcal{L}$  è già regolato da una convenzione di veracità in  $\mathcal{L}$  quale si applica agli imperativi e ai commissivi.

Si può obiettare che ci sono situazioni in cui il permesso viene richiesto, ma tale requisito non è stato prodotto da alcun imperativo né da alcuna promessa precedente. Di quali situazioni si tratti dipenderà dalle istituzioni di una particolare popolazione. Non dobbiamo aspettarci che tali situazioni siano trattabili puramente in termini di convenzioni di linguaggio. Forse non sono trattabili nei termini di nessuna convenzione. Ma possiamo concepirle come sottoprodotti dell'uso dei permessi in quanto neutralizzatori degli imperativi e dei commissivi. Possiamo dire che nella popolazione vale la regola o la convenzione di agire *come se* fossero stati emessi certi imperativi o certi commissivi: imperativi o commissivi con condizioni di verità che permettono la neutralizzazione.

Il nostro esame della veracità applicata ai modi non indicativi si ferma qui. Ci sono certo altri modi da esaminare, ma io non ho niente da dire a questo proposito e non credo che essi siano molto importanti ai fini della nostra comprensione del linguaggio.

Possiamo facilmente estendere il nostro esame a una lingua possibile  $\mathcal{L}$  con enunciati indicativi non ambigui, ai quali vengano assegnate interpretazioni diverse in diverse occasioni possibili della loro emissione. Di fatto, in italiano c'è carenza di enunciati eterni non indicativi, per cui abbiamo già dovuto far uso di enunciati indicativi per illustrare varie questioni. Cogliamo l'occasione per ricapitolare la formulazione di una convenzione di veracità in  $\mathcal{L}$ , ma questa volta esibendo dappertutto la dipendenza, tanto dei modi quanto delle condizioni di verità, dalle occasioni di emissione. Sia  $P$  una popolazione in cui vale la convenzione; sia  $x$  un membro qualunque di  $P$  (o quasi).

Se  $\sigma$  è un enunciato di  $\mathcal{L}$  che sarebbe indicativo in  $\mathcal{L}$  nell'occasione  $o$  della sua emissione da parte di  $x$ , indirizzata a un uditorio in  $P$ , allora

$x$  cerca di accertarsi del fatto che emette  $\sigma$  in  $o$  solo a condizione che  $\sigma$  sia vero in  $\mathcal{L}$  in  $o$ .

Se  $\sigma$  è un enunciato di  $\mathcal{L}$  che era imperativo in  $\mathcal{L}$  nell'occasione  $o$  della sua emissione da parte di un membro  $y$  di  $P$ , indirizzata a  $x$ ; se  $y$  e  $x$  hanno - ed è conoscenza comune tra loro che hanno - un interesse comune a rendere possibile a  $y$  il controllo delle azioni di  $x$ , entro una certa sfera, per mezzo dell'emissione di espressioni verbali; se in quella sfera c'è qualche azione per mezzo della quale  $x$  potrebbe cercare di rendere  $\sigma$  vero in  $\mathcal{L}$  in  $o$ ; allora  $x$  cerca di agire in modo tale che  $\sigma$  fosse vero in  $\mathcal{L}$  in  $o$ .

Se  $\sigma$  è un enunciato di  $\mathcal{L}$  che era commissivo in  $\mathcal{L}$  nell'occasione  $o$  della sua emissione da parte di  $x$ , indirizzata a un uditorio in  $P$ , allora  $x$  cerca di agire in modo tale che  $\sigma$  fosse vero in  $\mathcal{L}$  in  $o$ .

Come volevamo, le domande e i permissivi non sono stati menzionati separatamente: le domande perché sono annoverate tra gli imperativi, i permissivi perché rientrano nelle condizioni di verità di alcuni imperativi e commissivi.

Non ho ancora detto che cos'è la veracità rispetto agli enunciati ambigui di  $\mathcal{L}$ . Forse non è necessario affrontare il problema. Può darsi che basti dire semplicemente che  $\mathcal{L}$  è la lingua attuale di una popolazione  $P$  solo se in  $P$  c'è una regolarità convenzionale di veracità rispetto a quegli enunciati di  $\mathcal{L}$  che sono non ambigui in  $\mathcal{L}$  nelle loro occasioni di emissione. Questa convenzione potrebbe essere la conseguenza di una più generale convenzione di veracità in  $\mathcal{L}$  rispetto a tutti gli enunciati di  $\mathcal{L}$ , ambigui o no; ma potremmo dare una spiegazione della convenzione più limitata senza sapere che cos'è la veracità in  $\mathcal{L}$  rispetto agli enunciati ambigui.

Alternativamente, potremmo cercare di descrivere un livello minimo di veracità in  $\mathcal{L}$  rispetto agli enunciati ambigui. Prendiamo un enunciato  $\sigma$  al quale siano assegnate da  $\mathcal{L}$  molteplici interpretazioni in un'occasione  $o$  della sua emissione. Si può essere veraci in modo minimale in  $\mathcal{L}$  rispetto a  $\sigma$  in  $o$  prendendo una qualunque di quelle interpretazioni e facendo quello che si dovrebbe fare per essere veraci in  $\mathcal{L}$  se quella interpretazione fosse l'unica assegnata da  $\mathcal{L}$  a  $\sigma$  in  $o$ . Un bugiardo è verace in questo modo minimale se, sapendo che Paolo ha una predilezione per il frutto del pesco, dice "A Paolo piace la pesca" durante una conversazione sugli sport preferiti da Paolo.

Per descrivere un livello superiore di veracità dovremmo menzionare i metodi che applichiamo effettivamente per risolvere l'ambiguità nella pratica discorsiva. Penso che questi metodi dipendano dalle opinioni che ci facciamo sui nostri reciproci intenti discorsivi. Per esempio, possiamo igno-

rare un'interpretazione assegnata da  $\mathcal{L}$  a un enunciato  $\sigma$  nell'occasione  $o$  se quella interpretazione, ma non qualche altra, è discorsivamente inutile: se tra i partecipanti alla conversazione nell'occasione  $o$  è conoscenza comune che, se quella fosse l'unica interpretazione assegnata da  $\mathcal{L}$  a  $\sigma$  in  $o$ , l'emissione di  $\sigma$  in  $o$  non potrebbe servire ad alcun intento discorsivo dell'emittente. Possiamo ignorare un'interpretazione di un enunciato nella quale è conoscenza comune che l'enunciato è manifestamente falso. Possiamo ignorare un'interpretazione di un enunciato se è conoscenza comune che né l'emittente né alcun altro partecipante alla conversazione si accorgerà che l'enunciato ha quella interpretazione. Questi sono soltanto i più grossolani dei nostri metodi per risolvere l'ambiguità. Tuttavia, ho alcune esitazioni a suggerire che essi facciano parte del contenuto della nostra convenzione di veracità in una lingua. Sembra meglio concepirli come risultati dell'esercizio del buon senso in presenza di una convenzione di linguaggio. Ad ogni modo, non possiamo inserire i nostri metodi di risoluzione dell'ambiguità in una definizione di convenzione di veracità in  $\mathcal{L}$  finché non abbiamo compreso meglio la natura dei nostri metodi di risoluzione dell'ambiguità.

Ci accontenteremo quindi di concepire la veracità in una lingua possibile  $\mathcal{L}$  come veracità in  $\mathcal{L}$  rispetto agli enunciati non ambigui dei modi in questione, o eventualmente anche come veracità minimale in  $\mathcal{L}$  rispetto agli enunciati ambigui di quei modi.

Finora mi sono limitato a sostenere che, di fatto, gli utenti della lingua partecipano a convenzioni di veracità. Per verificare la mia asserzione dobbiamo semplicemente far ricorso alla nostra conoscenza di ciò che succede quando persone normali usano una lingua in modo normale. Possiamo immaginare cosa significa che una lingua possibile  $\mathcal{L}$  sia la lingua attuale di una popolazione  $P$  più o meno come la nostra, in circostanze più o meno come le nostre. Quando ce lo immaginiamo scopriamo secondo me che, sulla base delle mie definizioni, prevale in  $P$  una convenzione di essere veraci in  $\mathcal{L}$ : una convenzione sostenuta dall'interesse dei membri di  $P$  per la comunicazione, per la possibilità di controllare, in una certa misura, le proprie reciproche credenze e azioni mediante la produzione di suoni e di caratteri. Abbiamo esaminato questa asserzione prima in relazione alle lingue di segnalazione verbale; poi in relazione a lingue semplificate dalla rimozione dei modi non indicativi, dell'indicalità e dell'ambiguità; e infine in relazione a lingue in cui erano state reintrodotte queste complicazioni.

Se è vero che le convenzioni di veracità sono una caratteristica dell'uso normale della lingua quale noi lo conosciamo, e se - come suppongo - ne sono una caratteristica importante, allora potrebbe essere ragionevole farvi ricorso per definire cosa significa che una lingua possibile è la lingua attuale

di una popolazione  $P$ , Potremmo cioè adottare la definizione seguente:

$\mathcal{L}$  è una lingua attuale di  $P$  se e solo se in  $P$  prevale una convenzione di veracità in  $\mathcal{L}$ , consolidata da un interesse per la comunicazione.

Secondo la nostra precedente discussione sulla segnalazione verbale, questa definizione sarebbe adeguata nel caso particolare delle lingue di segnalazione verbale; la mia ipotesi è che sia adeguata anche in generale. Avanzo cioè l'ipotesi che essa si adatti all'uso comune nei casi chiari, e che tracci un'opportuna linea discriminante tra i casi non chiari. Il test per la definizione consisterebbe nel cercare di immaginare casi chiari in cui essa non concordasse con la nostra inclinazione ad affermare o a negare che  $\mathcal{L}$  è una lingua attuale di  $P$ ; io non sono riuscito a immaginarne neppure uno.

È possibile sostenere, per mezzo di controesempi, che la definizione non fornisce una condizione necessaria affinché  $\mathcal{L}$  sia una lingua attuale di  $P$ . Possiamo inventare molti esempi bizzarri in cui siamo inclini ad affermare che  $\mathcal{L}$  è una lingua attuale di  $P$  sebbene in  $P$  non ci sia alcuna convenzione di veracità in  $\mathcal{L}$ . Possiamo immaginare una popolazione di mentitori inveterati, o di persone ciascuna delle quali sospetta che gli altri siano mentitori inveterati, o di persone che usano la propria lingua solo per raccontare favole, o di creature istintive che sono incapaci di usare una lingua diversa da quella che è connaturata ad esse. Ma nessuno di questi bizzarri controesempi è convincente, perché non appena ci rendiamo conto di quanto sono eccezionali - di quanto sono lontani dall'uso della lingua quale noi lo conosciamo -, non siamo più disposti a classificarli come casi *chiari* nell'uso comune. Ma, se sono casi non chiari, siamo liberi di trattarli nel modo che troviamo più opportuno. Naturalmente, possiamo ammettere che siano casi in cui una lingua è, in un'accezione estesa del termine, la lingua attuale di una popolazione; ma questo equivale semplicemente a dire che essi hanno delle analogie importanti con i casi in cui la condizione espressa dalla definizione è soddisfatta.

## 5.5 Semantica in una popolazione

Nella misura in cui abbiamo formulato delle condizioni necessarie e sufficienti adeguate affinché una lingua sia la lingua attuale di una popolazione, siamo in grado di definire alcune relazioni semantiche tra espressioni verbali e popolazioni di utenti linguistici. Tali relazioni dipendono in modo ovvio dalle corrispondenti relazioni semantiche tra espressioni verbali e lingue possibili, grazie alla relazione che abbiamo esaminato tra lingue e popolazioni. In

generale, avremo di nuovo relazioni a quattro argomenti: questa volta, tra una espressione verbale  $\sigma$ , una popolazione  $P$ , un'occasione possibile  $o$  di emissione di  $\sigma$  e un'interpretazione  $\langle \mu, \tau \rangle$ .

$\sigma$  riceve da  $P$  in  $o$  l'interpretazione  $\langle \mu, \tau \rangle$  se e solo se esiste una lingua possibile  $\mathcal{L}$  tale che  $\mathcal{L}$  è una lingua attuale della popolazione  $P$  e  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$ .

Se  $\sigma$  è *eterno* in  $P$  - cioè riceve da  $P$  lo stesso insieme di interpretazioni in ogni occasione possibile della sua emissione -, possiamo omettere la menzione dell'occasione  $o$ . Se  $\sigma$  è *non ambiguo* in  $o$  - cioè riceve da  $P$  una singola interpretazione in  $o$  -, possiamo omettere la menzione dell'interpretazione  $\langle \mu, \tau \rangle$ .

$\sigma$  è *indicativo* in  $P$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se esiste una lingua possibile  $\mathcal{L}$  tale che:  $\mathcal{L}$  è una lingua attuale della popolazione  $P$ ;  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$ , e  $\mu$  è 0. (Analogamente per gli altri modi citati.)

$\sigma$  è *vero* in  $P$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se esiste una lingua possibile  $\mathcal{L}$  tale che:  $\mathcal{L}$  è una lingua attuale della popolazione  $P$ ;  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$ , e la condizione di verità  $\tau$  vale nel - contiene il - mondo possibile  $w$  nel quale è situata l'occasione possibile  $o$  di emissione di  $\sigma$ . (Analogamente per la falsità.)

$\sigma$  è *analitico* in  $P$  in  $o$  nell'interpretazione  $\langle \mu, \tau \rangle$  se e solo se esiste una lingua possibile  $\mathcal{L}$  tale che:  $\mathcal{L}$  è una lingua attuale della popolazione  $P$ ;  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un insieme di interpretazioni contenente  $\langle \mu, \tau \rangle$ , e la condizione di verità  $\tau$  vale in ogni mondo possibile. (Analogamente per la contraddittorietà e la sinteticità.)

Semplificando per il caso in cui  $\sigma$  è sia eterno che non ambiguo in  $P$ , otteniamo definizioni come le seguenti:

$\sigma$  è *vero* in  $P$  nel mondo  $w$  se solo se esiste una lingua possibile  $\mathcal{L}$  tale che:  $\mathcal{L}$  è una lingua attuale della popolazione  $P$ ;  $\mathcal{L}$  assegna a  $\sigma$ , in ogni occasione possibile della sua emissione, un'unica interpretazione fissa  $\langle \mu, \tau \rangle$  la cui condizione di verità  $\tau$  vale nel mondo possibile  $w$ .

$\sigma$  è *vero* in  $P$  se e solo se esiste una lingua possibile  $\mathcal{L}$  tale che:  $\mathcal{L}$

è una lingua attuale della popolazione  $P$ ;  $\mathcal{L}$  assegna a  $\sigma$ , in ogni occasione possibile della sua emissione, un'unica interpretazione fissa  $\langle \mu, \tau \rangle$  la cui condizione di verità  $\tau$  vale nel mondo attuale.

$\sigma$  è *analitico* in  $P$  se e solo se esiste una lingua possibile  $\mathcal{L}$  tale che:  $\mathcal{L}$  è una lingua attuale della popolazione  $P$ ;  $\mathcal{L}$  assegna a  $\sigma$ , in ogni occasione possibile della sua emissione, un'unica interpretazione fissa  $\langle \mu, \tau \rangle$  la cui condizione di verità  $\tau$  vale in ogni mondo possibile.

Possiamo semplificare i definenda in un modo diverso, conservando la menzione dell'occasione  $o$  di emissione di  $\sigma$ , ma non menzionando la popolazione  $P$ . Infatti, esaminando  $o$  possiamo identificare un comunicante e l'uditorio cui si rivolge, e ricercare una lingua attuale della popolazione cui essi appartengono. Otteniamo così delle relazioni semantiche tra un'espressione verbale a  $\sigma$  un'occasione  $o$  della sua emissione (e, se dobbiamo provvedere all'ambiguità, un'interpretazione). Non occorre menzionare né una popolazione né una lingua. Illustrerò i definenda semplificati senza occuparmi dell'ambiguità.

$\sigma$  è *vero* in  $o$  se e solo se esistono una lingua possibile  $\mathcal{L}$  e una popolazione  $P$  tali che: il comunicante e l'uditorio cui egli si rivolge in  $o$  appartengono a  $P$ ;  $\mathcal{L}$  è una lingua attuale di  $P$ ;  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un'unica interpretazione  $\langle \mu, \tau \rangle$  la cui condizione di verità  $\tau$  vale nel mondo possibile  $w$  in cui è situata  $o$ .

$\sigma$  è *analitico* in  $o$  se e solo se esistono una lingua possibile  $\mathcal{L}$  e una popolazione  $P$  tali che: il comunicante e l'uditorio cui egli si rivolge in  $o$  appartengono a  $P$ ;  $\mathcal{L}$  è una lingua attuale di  $P$ ;  $\mathcal{L}$  assegna a  $\langle \sigma, o \rangle$  un'unica interpretazione  $\langle \mu, \tau \rangle$  la cui condizione di verità  $\tau$  vale in ogni mondo possibile.

Questo per le interpretazioni date da una popolazione a quelle espressioni verbali che sono gli enunciati della sua lingua. Che dire dei costituenti che non sono costituenti d'enunciato in una grammatica di quella lingua? Anch'essi hanno delle interpretazioni. (Trascureremo l'indicalità, l'ambiguità e l'anomalia, e fingeremo che abbiano delle singole interpretazioni fisse.) La grammatica assegna delle interpretazioni a tutti i suoi costituenti, sia direttamente che per mezzo delle sue operazioni di proiezione. Ma noi vorremmo dire, per esempio, che a una parola viene assegnata un'interpretazione da una popolazione di utenti linguistici. I gallesi assegnano alla parola "gwyn" una certa interpretazione: forse si tratta della funzione che assegna a ogni mondo possibile l'insieme delle cose bianche che vi sono contenute. Nel di-

re questo, non abbiamo menzionato né una lingua né una grammatica. Ci occorre una relazione a tre posti: il costituente  $\xi$  riceve dalla popolazione  $P$  l'interpretazione  $\iota$ .

Come possiamo definire questa relazione? Presumibilmente menzionando una lingua  $\mathcal{L}$  e una grammatica  $\Gamma$  che colleghi la popolazione  $P$  con l'assegnazione di  $\iota$  a  $\xi$ .  $\Gamma$  assegna  $\iota$  a  $\xi$ ;  $\mathcal{L}$  è data da  $\Gamma$ ; e in  $P$  vale una convenzione di veracità in  $\mathcal{L}$ . Sfortunatamente  $\mathcal{L}$  non avrà una grammatica sola. Grammatiche di  $\mathcal{L}$  diverse interpreteranno diversamente i costituenti; può darsi che una di esse assegni  $\iota$  a  $\xi$  e un'altra no. Può darsi addirittura che grammatiche di  $\mathcal{L}$  diverse analizzino in modo diverso gli enunciati in costituenti; può darsi che  $\xi$  sia un costituente in una grammatica di  $\mathcal{L}$  ma non in un'altra. Queste differenze tra grammatiche di  $\mathcal{L}$  si annullano; le diverse grammatiche danno gli stessi enunciati, con le stesse interpretazioni. Data  $P$ , scegliamo  $\mathcal{L}$  andando alla ricerca di una convenzione di veracità; ma, data  $\mathcal{L}$  come possiamo scegliere  $\Gamma$ ? Le convenzioni di veracità riguardano gli enunciati interi e lasciano indeterminate le interpretazioni delle parti degli enunciati. Dovremmo forse andare alla ricerca di convenzioni di qualche altra specie, ma non riesco a immaginare quale potrebbe essere il contenuto di una simile convenzione. Non potrebbe essere semplicemente la convenzione di adottare una grammatica così- e- così, o la convenzione di assegnare interpretazioni così- e- così a costituenti così- e- così. Le convenzioni sono regolarità d'azione, e non esiste un'azione come quella che consiste nell'adottare una grammatica o nell'assegnare una interpretazione (o, se esiste, non si verifica nell'uso normale della lingua).

Alcune grammatiche di  $\mathcal{L}$  sono semplici, naturali, ragionevoli, facili, buone; altre sono complicate, artificiose, macchinose, difficili, cattive. Forse  $\Gamma$  dovrebbe essere la *migliore* grammatica di  $\mathcal{L}$ , in base a qualche opportuno metodo di valutazione delle grammatiche alternative. Possiamo definire una relazione a quattro argomenti tra un costituente  $\xi$ , una popolazione  $P$ , un'interpretazione  $\iota$  e un metodo di valutazione  $M$ :

$\xi$  riceve da  $P$  l'interpretazione  $\iota$  (in base a  $M$ ) se e solo se esistono una lingua possibile  $\mathcal{L}$  e una grammatica  $\Gamma$  di  $\mathcal{L}$ , tali che:  $\mathcal{L}$  è una lingua attuale della popolazione  $P$ ;  $\xi$  è un costituente in  $\Gamma$  e  $\Gamma$  assegna  $\iota$  a  $\xi$ ;  $\Gamma$  è la migliore grammatica di  $\mathcal{L}$  in base al metodo di valutazione  $M$ .

Se riusciamo a scoprire un metodo di valutazione che sia l'unico a meritare una condizione privilegiata, possiamo eliminare la sgradita relativizzazione a  $M$ . Altrimenti, temo che la nozione che ci occorre sia improponibile: non

esiste una cosa come *la* interpretazione assegnata da *P* a  $\xi$ , ma ci sono soltanto le varie interpretazioni assegnate da *P* a  $\xi$  in base ai vari metodi alternativi di valutare le grammatiche.

Perché un qualche metodo di valutazione delle grammatiche potrebbe meritare una condizione privilegiata? Forse perché, come ha ipotizzato Chomsky, esso è il solo che entra nella spiegazione psicologica della competenza linguistica.<sup>11</sup> Supponiamo che, ogniqualvolta qualcuno acquisisce la capacità e la tendenza a essere verace e ad aspettarsi la veracità in una lingua, egli si costruisca una specie di rappresentazione interiore inconscia di una grammatica di quella lingua; e supponiamo che, in qualunque stadio dell'acquisizione del linguaggio da parte di un bambino, la grammatica che egli si rappresenta interiormente sia la grammatica che *meglio* si adatta, in base a un certo metodo fisso *M* di valutazione, all'uso della lingua che ha osservato attorno a sé. È quindi *M*, assieme all'osservazione, che determina quale grammatica un bambino si rappresenterà interiormente; e la grammatica che i parlanti nativi di una lingua  $\mathcal{L}$  si rappresentano interiormente sarà la grammatica di  $\mathcal{L}$  che è migliore in base a *M*.

Questa ipotesi psicolinguistica è ancora in uno stadio congetturale. Io posso dire soltanto che se essa, o qualcosa di simile, è vera, allora c'è un senso preciso in cui si può dire che un costituente riceve un'interpretazione da una popolazione. Altrimenti, dovremo cercare altrove un metodo di valutazione privilegiato, oppure dovremo rinunciare. Rinunciare significherebbe accettare la tesi di Quine dell'imperscrutabilità del riferimento: indipendentemente da quanto sappiamo dei membri della popolazione *P*, non disponiamo di alcun modo oggettivamente determinato di interpretare parti del loro discorso più piccole degli enunciati.<sup>12</sup>

Spero che a questo punto sia chiaro cosa significa che un'espressione verbale *a* è analitica in una popolazione *P*. Rimane aperto il problema di sapere se una qualunque espressione verbale *sia* mai analitica in una popolazione *P*. L'analiticità, quale è stata descritta finora, potrebbe essere una prerogativa perfettamente intelligibile ma non posseduta da alcunché. Analogamente, possiamo sapere cosa significa che una lingua possibile  $\mathcal{L}$  è la lingua attuale di una popolazione; ma non sappiamo se ciò si verifica. E io dubito fortemente che non si verifichi.

Si dice spesso che l'analiticità non è chiara. (Dire questo è molto diverso dal dire che l'analiticità, chiara o non chiara, è inintelligibile.) In qualunque

---

<sup>11</sup> *Aspects of the Theory of Syntax*, cit., cap. I.

<sup>12</sup> *Word and Object*, pp. 68-79 [trad. it. cit., pp. 89-99. (N.d.T.)]; "Ontological Relativity: The Dewey Lectures 1968", *Journal of Philosophy*, 65, 1968, pp. 185-198.

popolazione, per esempio nella nostra, la maggior parte degli enunciati è chiaramente sintetica, alcuni enunciati insignificanti sono abbastanza chiaramente analitici, e tutto ciò che è interessante da un punto di vista filosofico e scientifico sembra in qualche modo stare in mezzo. Per esempio, sembra che, in qualche punto, ci sia dell'analiticità nei principi fondamentali della dinamica. Ma dove? Nella conservazione della quantità di moto? Nel principio di azione e reazione? Nel principio secondo cui la forza è uguale alla massa per l'accelerazione? In qualche altro punto? Non sappiamo dirlo. Ciascuno dei principi menzionati sembra, in qualche modo, parzialmente analitico. Ma com'è possibile l'analiticità parziale?

La risposta a questa domanda non risiede in modo primario nel fatto che le nostre convenzioni di linguaggio, per quanto siano tali, non sono convenzioni al grado più elevato. Non sembra probabile che un grado più elevato di convenzionalità (ossia un numero minore di casi di non verità e di eccezioni di vario genere, e una maggiore fiducia nella verità altrui) renderebbe più chiara l'analiticità. Né quella risposta risiede in modo primario nel fatto che non è chiaro esattamente quali mondi possibili ci sono, sebbene ciò non sia chiaro. (Per esempio, io non ho idea se un qualsiasi mondo possibile è a cinque dimensioni.) Dato che l'analiticità è verità - nella lingua di una popolazione - in ogni mondo possibile, l'incertezza riguardo alla possibilità dei mondi potrebbe certamente avere come riflesso l'incertezza riguardo all'analiticità degli enunciati. Ma questa non può essere la spiegazione completa dell'analiticità non chiara. A volte non siamo in grado di dire se un enunciato è analitico - come nel caso, per esempio, di uno dei principi della dinamica - perché non siamo in grado di dire se è vero nella nostra lingua in qualche mondo ipotetico che è chiaramente possibile.

Penso che dovremmo concludere che una convenzione di verità in una singola lingua possibile è un caso limite - mai pienamente realizzato - di qualcos'altro: una convenzione di verità in una lingua qualunque da noi scelta entro una famiglia ristretta di lingue possibili molto simili. Le lingue di tale famiglia hanno esattamente gli stessi enunciati e assegnano a essi insieme corrispondenti di interpretazione; ma a volte ci sono lievi differenze nelle corrispondenti condizioni di verità. È raro che queste differenze influiscano su mondi abbastanza vicini a quello attuale da essere compatibili con la maggior parte delle nostre credenze ordinarie. Ma, via via che passiamo a mondi possibili sempre più bizzarri, un numero sempre maggiore dei nostri enunciati si rivela vero in alcune lingue della nostra famiglia e falso in altre. Così può accadere che un enunciato sia analitico in alcune lingue della nostra famiglia ma falso (in mondi sufficientemente remoti) in altre. Per noi quell'enunciato sarebbe *parzialmente* o *non chiaramente* analitico. La no-

stra lingua attuale è come un ibrido in cui echeggiano le lingue possibili che lo costituiscono. L'analiticità è chiara in qualunque di esse, ma può darsi che esse non concordino.

Anche se non spiegasse il problema dell'analiticità non chiara, l'ipotesi che le nostre convenzioni di linguaggio ci vincolino a famiglie, e non a singole lingue possibili, sarebbe plausibile per altri motivi. La specie di convenzione che ho in mente è la seguente: quasi ciascuno, quasi sempre, è verace almeno in alcune lingue della famiglia, ma non necessariamente esse sono le stesse per ciascuno, o per una persona in momenti diversi. Avere una convenzione come questa è per noi meglio che avere una convenzione di veracità in una singola lingua. La standardizzazione a fini comunicativi è una buona cosa, ma non è essenziale. Se la famiglia è sufficientemente ristretta, la comunicazione non risulta compromessa. Di solito noi non parliamo di mondi possibili remoti in cui le condizioni di verità divergono; di solito, quindi, la veracità in una lingua qualunque della famiglia coincide con la veracità in ogni lingua della famiglia. Quando si manifesta una divergenza che ci disorienta, possiamo salvare la comunicazione - ammesso che abbiamo un po' di buona volontà - andando alla ricerca di enunciati le cui condizioni di verità non divergano in modo fastidioso e formulando convenzioni temporanee più restrittive della nostra convenzione permanente. In cambio del fatto che abbiamo accettato di correre qualche rischio per quanto riguarda la possibilità della comunicazione, otteniamo un duplice vantaggio.

Da un lato, le differenti lingue della famiglia possono avere pregi e difetti diversi, per cui possono essere diversamente intonate alle opinioni, ai gusti e agli intenti discorsivi individuali. Se ciascuno può scegliere entro la famiglia, le preferenze incompatibili riguardo alle lingue saranno tutte soddisfatte. Inoltre, non limitandoci a una singola lingua, evitiamo il rischio di affidarci esclusivamente a una singola lingua che in seguito può rivelarsi inopportuna alla luce di nuove scoperte e di nuove teorie; e ci garantiamo una certa flessibilità senza che sia necessario cambiare la convenzione.

Ma - ciò che è più importante - quanto meno la nostra convenzione permanente è restrittiva, tanto minore è la quantità di esperienza necessaria a identificarla, a impadronirsene e a diventare partecipi di essa. Un bambino deve effettuare estrapolazioni da un campione limitato di uso linguistico, una parte del quale viola addirittura la convenzione. Supponiamo che tutti coloro che gli stanno intorno siano veraci, per tutto il tempo, esattamente nelle stesse lingue possibili. Potrebbe accadere che il bambino debba aspettare ancora a lungo prima di osservare una quantità di discorso sufficiente a permettergli di identificare la lingua. Infatti, se due lingue differiscono soltanto riguardo ai valori di verità degli enunciati in mondi possibili remo-

ti, la differenza si manifesterà nel giro di pochissime conversazioni. Bisogna riconoscere che il bambino è aiutato dalla sua tendenza a ignorare quasi tutte le estrapolazioni compatibili con i suoi dati, altrimenti non riuscirebbe a conseguire alcun risultato. Ma, a condizione che più di una lingua della famiglia sia seriamente candidata - e io ritengo che ciò sia dimostrato dall'esistenza dell'analiticità non chiara -, egli ha bisogno di più informazioni per identificare una lingua entro la famiglia che per identificare la famiglia. Avanzo quindi l'ipotesi che una convenzione di veracità in una singola lingua possibile non potrebbe consolidarsi. Verrebbe appresa imperfettamente; essendo stata appresa imperfettamente, verrebbe violata di frequente; essendo stata violata frequentemente, verrebbe appresa ancora più imperfettamente.

# Conclusione

Nell'introduzione mi sono assunto il compito di sostenere il luogo comune secondo cui esistono delle convenzioni di linguaggio. Così ho fatto, non so se in modo convincente. Prima, e a lungo, ho tentato di dire in generale che cos'è una convenzione; poi, per sommi capi, ho cercato di dire quali specie di convenzione potrebbero essere convenzioni di linguaggio di una popolazione. Così ho cercato di rispondere alla sfida scettica di Quine e White.

Ma il loro scetticismo riguardo alle convenzioni di linguaggio è solo un aspetto di uno scetticismo di più ampia portata: quello riguardo all'analiticità, alla contraddittorietà, alla sinteticità, all'implicazione logica, alla sinonimia (nel senso di implicazione logica reciproca) e a tutto il resto. Com'è ormai noto, queste nozioni sono facilmente interdefinibili. Se un membro qualunque della catena ha senso, hanno senso tutti i membri. Io ho cercato di dare un senso all'analiticità. O, piuttosto, ho cercato di dare un senso a due diverse analiticità. Nel capitolo 5.3 ho definito l'analiticità in una lingua possibile; nel capitolo 5.5 ho definito l'analiticità in una popolazione  $P$  di utenti linguistici. Quest'ultima analiticità dipendeva dalla prima e dalla relazione convenzionale tra una lingua possibile e una popolazione. Nell'ipotesi che sia riuscito a salvare il luogo comune secondo il quale il linguaggio è convenzionale, sono riuscito, di conseguenza, a salvare l'analiticità e le nozioni gemelle? Io credo di sì.

Ci sono due tipi di teoria semantica, a volte erroneamente considerati come tentativi contrastanti di svolgere la stessa funzione, consistente nell'analizzare proprietà semantiche come la verità, l'analiticità e simili. Non c'è alcuna funzione del genere da svolgere, perché tali proprietà non esistono. Esistono invece coppie di *relazioni* semantiche: verità in una lingua  $\mathcal{L}$ , verità in una popolazione  $P$ ; analiticità in  $\mathcal{L}$ , analiticità in  $P$ ; e così via.

Un tipo di semantica analizza la verità, l'analiticità e le nozioni affini in relazione a lingue possibili interpretate, facendo astrazione dagli utenti di queste ultime. È il tipo di semantica sviluppata da Frege, da Tarski e (per lo più) da Carnap. Di solito questo tipo di semantica era afflitta

dall'incapacità di distinguere tra espressioni coestensive con senso diverso, e dall'incapacità di maneggiare gli aspetti indicali del linguaggio; ma a queste mancanze si è ovviato di recente applicando l'idea che le estensioni vanno assegnate relativamente a un mondo possibile e relativamente al momento dell'emissione, al parlante ecc.

L'altro tipo di semantica analizza la verità, l'analiticità e le nozioni affini in relazione a un agente o a una popolazione di agenti. Si tratta di un orientamento sviluppato dal secondo Wittgenstein, da Grice, da Skinner, da Quine, da Morris, da Ziff e (a volte) da Carnap; a sua volta, esso è afflitto di sovente dalla riluttanza a prendere in esame più di un'espressione verbale per volta.

Elude il problema chi afferma che il primo tipo di semantica si occupa delle lingue artificiali, mentre il secondo tipo si occupa delle lingue naturali. Entrambe le semantiche si sono occupate di entrambi gli argomenti; entrambe sono state applicate con più successo a frammenti semplificati delle lingue naturali che a lingue naturali complete, com'era prevedibile. Il gioco linguistico del muratore e dell'aiutante<sup>13</sup> è simile a una lingua naturale quasi quanto un calcolo del primo ordine.

Passiamo ora ad esaminare una parte delle critiche mosse da Quine all'analiticità. Quine lamenta che le spiegazioni dell'analiticità fornite da Carnap e dagli altri siano prive di utilità. Egli prende in esame la proposta secondo la quale un'asserzione è analitica in  $\mathcal{L}$  se e solo se è vera in virtù delle regole semantiche di  $\mathcal{L}$  (e di queste soltanto); e secondo la quale la lingua  $\mathcal{L}$  - almeno se è una lingua artificiale - è una coppia ordinata il cui secondo componente è un insieme di regole semantiche che specificano delle condizioni di verità. Quine risponde così:

“Ma, per la stessa ragione e più semplicemente, potremmo interpretare una lingua artificiale  $L$  direttamente come una coppia ordinata di cui il secondo componente sia la classe delle sue asserzioni analitiche; e allora le asserzioni analitiche di  $L$  divengono determinabili semplicemente come le asserzioni che fan parte del secondo componente di  $L$ . O, ancor meglio, potremmo finirla di arrampicarci sugli specchi.

Fare appello a lingue ipotetiche artificialmente elementari potrebbe essere un modo utile per chiarire l'analiticità se i fattori mentali o comportamentali o culturali di rilievo per l'analiticità (quali che possano essere) fossero in qualche modo abbozzati in quel modello

---

<sup>13</sup>Cfr. L. Wittgenstein, *Ricerche filosofiche*, trad. it. cit., §2 sgg. (N.d.T.)

semplificato. Ma un modello che prende l'analiticità puramente come una caratteristica non riducibile non è certo fatto per gettar luce sul problema dell'analiticità.”<sup>14</sup>

Cosa succede? Non importa che si supponga che l'oppositore carnapiano di Quine stia parlando di lingue artificialmente semplificate. Mettiamogli in bocca la più audace proposta secondo cui qualunque lingua possibile può essere rappresentata da una coppia ordinata il cui secondo componente è un insieme di regole semantiche di verità. (Ciò che è peculiare di una lingua artificiale è semplicemente il fatto che un filosofo può sperare di esibire quelle regole senza troppa fatica.) Non è esattamente questa la rappresentazione che abbiamo scelto per le nostre lingue possibili, ma è abbastanza simile. Tutto ciò è arrampicarsi sugli specchi, secondo Quine; non sarebbe invece arrampicarsi sugli specchi menzionare i fattori mentali, comportamentali o culturali.

Quine vuole sapere della analiticità nella lingua di una data popolazione.<sup>15</sup> Altrimenti non tocca a lui ammettere che qualunque informazione relativa alla popolazione potrebbe essere pertinente. Egli ha perfettamente ragione a negare che la sua domanda abbia ricevuto una risposta; il suo oppositore, infatti, gli sta parlando dell'analiticità in una lingua possibile, indipendentemente dall'uso che qualunque popolazione fa di quella lingua. Forse l'oppositore supponeva che Quine sapesse cosa significa che una lingua possibile sia la lingua di una popolazione. E in un certo senso Quine lo sapeva: egli sa cosa significa che la lingua dei gallesi *sia* la lingua dei gallesi. Ma quello che Quine *non* conosceva - e che il suo oppositore carnapiano non gli ha spiegato<sup>16</sup> - era la relazione che deve sussistere tra i gallesi e una certa coppia ordinata (il cui secondo componente è un insieme di regole seman-

---

<sup>14</sup>“Two Dogmas of Empiricism”, pp. 35-36. [Cfr. trad. it. cit. pp. 34-35. (N.d.T.)]

<sup>15</sup>Ciò è coerente con “Two Dogmas of Empiricism”, p. 33: “Una asserzione *S*, si dice, è *analitica* per una lingua *L*, e il problema sta nel dare un senso a codesta relazione in generale, cioè per ‘*S*’ e ‘*L*’ variabili.” [Cfr. trad. it. cit., p. 32. (N.d.T.)] Il passo, infatti, può e deve essere interpretato nel senso che il problema è di dare un senso all'analiticità per una lingua *attuale* arbitraria, identificata non dalle sue proprietà semantiche ma dai suoi utenti.

<sup>16</sup>Il vero Carnap ha qualcosa da dire, in “Meaning and Synonymy in Natural Languages”, *Philosophical Studies*, 6 (1955), pp. 33-47. [Trad. it. di G. Usberti, “Significato e sinonimia nelle lingue naturali”, in *La struttura logica del linguaggio*, a cura di A. Bonomi, Milano, Bompiani, 1973. (N.d.T.)]. L'ipotesi che una data lingua possibile  $\mathcal{L}$  sia la lingua attuale di Karl ha delle conseguenze verificabili relative alle probabili risposte di Karl a domande del tipo “Che cosa diresti se...?” La difficoltà sta nel fatto che, per derivare queste conseguenze dall'ipotesi, dobbiamo prima capirla - cioè dobbiamo sapere che cosa vuoi dire che  $\mathcal{L}$  è la lingua attuale di Karl.

tiche) affinché quella coppia ordinata sia la lingua dei gallesi. Per spiegare questo a Quine, l'oppositore dovrà certamente menzionare i fattori mentali, comportamentali o culturali presenti tra i gallesi.

Sembra strano che Quine non si sia dimostrato scettico nei confronti della verità stessa. Egli ritiene che l'indagine di Tarski abbia chiarito la nozione di verità. Ma quello che Tarski ha chiarito è la nozione di verità in una data lingua possibile interpretata, non la nozione di verità nella lingua di una popolazione. Ammettiamo che "Gwyn yw eira" sia vero in una certa lingua possibile se e solo se la neve è bianca; ecco Morgan che dice "Gwyn yw eira": sta dicendo la verità? Il fatto che la lingua possibile in questione sia chiamata gallese perché è la lingua dei gallesi non spiega quello che intendiamo quando affermiamo che Morgan dice la verità nella sua lingua, né quello che intendiamo quando affermiamo semplicemente che Morgan dice la verità. Per spiegarlo, dovremmo parlare di fattori mentali, comportamentali e culturali.

Dapprima io ho fatto, più o meno, quello che faceva l'oppositore carnapiano di Quine: ho dato una spiegazione della natura di qualunque lingua possibile, in modo tale che sia facile identificare gli enunciati analitici di una data lingua possibile. Ma poi ho cercato di portare a termine il mio compito facendo quello che l'oppositore carnapiano non aveva fatto. Ho dato una spiegazione dei fattori mentali, comportamentali e culturali che operano tra i membri di una popolazione e determinano gli enunciati analitici della loro lingua determinando quale lingua possibile è la loro. Quindi ritengo di aver risposto alla richiesta di Quine: ho dato una spiegazione della specie di analiticità opportuna, cioè l'analiticità in relazione a una popolazione di utenti linguistici.

Tuttavia la mia risposta delude le speranze di alcuni filosofi. Spesso ho parlato liberamente di mondi possibili, e non vedo come avrei potuto eliminarli. Non è, in fondo, il vecchio circolo vizioso? Infatti, ogni modo conosciuto di spiegare i mondi possibili fa appello, in ultima analisi, all'analiticità. I mondi possibili sono modelli, o indici di modelli, o diagrammi di modelli (descrizioni di stato), o teorie dei modelli (resoconti completi) di qualche lingua sufficientemente ricca.<sup>17</sup> Ma, naturalmente, essi non so-

---

<sup>17</sup>Per esempio, i mondi possibili sono spiegati come modelli in Richard Montague, "Logical Necessity, Physical Necessity, Ethics, and Quantifiers", *Inquiry*, 3 (1960), pp. 259-269; come indici di modelli in Saul Kripke, "Semantical Considerations on Modal Logic" [Trad. it. cit. (N.J.T.)]; come descrizioni di stato in Carnap, "Modalities and Quantification", *Journal of Symbolic Logic*, 11 (1946), pp. 33-64; e come resoconti completi in Jeffrey, *Logic of Decision*, pp. 196-197.

no *tutti* i modelli della lingua, bensì solo quelli che soddisfano le verità analitiche!

Anche questo risultato ha più della spirale che del cerchio. Se all'inizio sapessimo soltanto che cos'è l'analiticità in una lingua particolare, potremmo concepire i nostri mondi possibili come modelli, o qualcos'altro, di quella lingua, e continuare a servirci di questi mondi possibili per spiegare l'analiticità in una lingua arbitraria. Sarebbe già un passo avanti.

Ma io direi piuttosto che l'idea stessa di rendere comprensibile la possibilità a vantaggio dell'analiticità era cattiva. Perché dovrebbero essere le convenzioni umane a creare e a distruggere i fatti riguardanti ciò che è possibile? Tutto quello che una convenzione umana può fare è scegliere un'espressione verbale piuttosto che un'altra per affidarle il privilegio della verità in virtù dei fatti riguardanti la possibilità dei mondi. Esattamente nello stesso modo, la convenzione umana sceglie un'espressione verbale piuttosto che un'altra per affidarle il privilegio della verità in virtù dei fatti riguardanti il tempo di domani. Volendo, possiamo dire che è per convenzione che non ci sono degli uomini sposati *correttamente chiamati* scapoli. Ma non possiamo dire che è per convenzione che non ci sono scapoli sposati, in questo o in qualunque altro mondo. Non potrebbero esserci.

Ad ogni modo, con quale diritto diciamo che i mondi possibili sono entità screditate, inadatte all'impiego filosofia) a meno che chiedano la redenzione alla filosofia del linguaggio? Non conosco accusa contro i possibili che non possa essere rivolta altrettanto giustamente contro gli insiemi. Eppure, poche coscienze filosofiche si fanno degli scrupoli con la teoria degli insiemi. Gli insiemi e i possibili tendono in modo analogo verso una ontologia affollata. Gli insiemi e i possibili pongono in modo analogo questioni per le quali non abbiamo risposta. Ci sono insiemi non numerabili più piccoli del continuo? Ci sono mondi possibili a cinque dimensioni? Io propongo di rimanere ugualmente tranquilli di fronte a questi misteri ugualmente misteriosi.

C'è un'accusa contro i possibili che non viene ripetuta contro gli insiemi; ma i possibili non sono colpevoli come viene loro imputato. Si dice che i possibili resistono all'individuazione. Quand'è che abbiamo un possibile in due mondi, e quando due possibili simili in due mondi? Una teoria drastica risolve facilmente la questione: nulla è mai in più di un mondo possibile. Questa soluzione estrema mostra che la difficoltà non è seria; forse soluzioni meno drastiche andrebbero anche meglio.<sup>18</sup>

I possibili si comportano abbastanza bene; quindi il problema di ammetterli o meno nelle nostre teorizzazioni metafisiche dipende dalla loro uti-

---

<sup>18</sup>Cfr. Lewis, "Counterpart Theory and Quantified Modal Logic".

lità. Io credo che qualunque tentativo equanime di farne uso nell'ontologia sistematica dimostrerà che essi sono di grande utilità.

L'analiticità è verità in tutti i mondi possibili. Ciò che è analitico per qualcuno dipende congiuntamente dai fatti relativi ai mondi possibili e dalla lingua che egli sta usando. La lingua che egli usa dipende dalle convenzioni delle quali è partecipe. E queste convenzioni sono regolarità di comportamento, consolidate dall'interesse per la coordinazione e dall'aspettativa che gli altri eseguano la propria parte.