

Bruno Cancellieri, ChatGPT et Claude.ai

Fondements de l'agentologie

Des entités aux agents

29 août 2026

Copyright © Bruno Cancellieri

« C'est un procrastinateur. » « Elle est envahissante. » « Je suis fait comme ça. » Ce sont des phrases qui semblent expliquer un comportement et qui en réalité ferment la discussion : elles attribuent une propriété stable, et d'une propriété stable il ne suit rien, sinon la possibilité de la constater de nouveau.

L'agentologie change les questions. Non pas ce qu'une chose est, mais qui agit, ce que chacun perçoit, quel objectif il poursuit, quelles figures il reconnaît entre lui et les autres, selon quelles logiques il décide, quelles actions il croit avoir à sa disposition. Ce sont des questions dont les réponses pourraient être autres qu'elles ne sont, et chacune indique un point où quelque chose peut être changé. Le même vocabulaire décrit une personne, une entreprise, une cellule et un programme informatique.

Il en découle trois choses utilisables même par qui n'a pas d'intérêt pour les appareils conceptuels. Lorsqu'on n'arrive pas à faire ce qu'on a décidé de faire, il n'y a pas de défaut de volonté : il y a deux parties de la même personne qui poursuivent des objectifs différents, et savoir comment elles sont faites dit où intervenir. Les émotions ne sont pas des états à subir mais des messages, avec un expéditeur, un objet et une demande : à une demande on peut répondre au lieu d'obéir. Et quand un rapport se détériore, presque jamais les deux n'y reconnaissent les mêmes figures : l'un croit rendre un service, l'autre reçoit un commandement, et aucun des deux ne ment.

Cet essai n'enseigne pas comment traiter les autres. Il invite à ne pas s'arrêter à la première explication qui vient à l'esprit, laquelle est d'ordinaire la plus commode et presque toujours inutile.

Sommaire

Note.....	5
Préface.....	6
Concepts fondamentaux.....	8
Agent	8
Conscience.....	8
Situation.....	8
Perception.....	8
Programme	9
Action	9
Axiomes	10
Définitions.....	11
Principe général et méthode agentologique	13
Questionnaire agentologique	13
Les émotions	14
Le sens des émotions.....	14
L'émotion comme message	14
La force du message : demandes et commandements.....	14
Attraction et répulsion.....	15
Ce que la conscience peut faire du message.....	15
Ce qui change, à les traiter ainsi.....	16
Les figures relationnelles	17
La situation relationnelle et les figures.....	17
La figure relationnelle comme entrée de la décision.....	17
La reconnaissance des figures relationnelles.....	18
Répertoire des figures relationnelles.....	19
Ce qu'il faut attendre de la liste.....	19
Sommaire des figures relationnelles.....	19
Les figures relationnelles quant aux objectifs	20

Les figures relationnelles quant aux programmes et à leur connaissance	21
Les figures relationnelles quant à la communication	21
Les figures relationnelles d'asymétrie.....	22
Les figures relationnelles quant à la valence et aux effets	23
La sélection.....	24
Le troisième agent	24
Le désalignement.....	25
Application aux relations sociales.....	27
Ce que dit la Psychologie de l'interagir	27
Ce qu'ajoute l'agentologie.....	27
La méta-interaction.....	28
Application à la psychothérapie et à l'autothérapie.....	30
Conflit intra-agent	30
Les niveaux du conflit	30
Pouvoir intra-agent et acrasie	31
Le changement thérapeutique comme apprentissage dirigé.....	31
La méthode appliquée à soi-même	31
Le problème spécifique de l'autothérapie	32
Avertissement sur l'usage de l'agentologie comme instrument thérapeutique	32
Un exemple d'application.....	33
Conclusion	35

Note

Cet essai est né d'un dialogue avec ChatGPT et Claude.ai, selon une méthode que j'ai adoptée pour d'autres travaux également : je propose un thème, je discute avec ces modèles de langage de son cadrage et de son contenu, et je fais écrire puis corriger les textes au moyen d'une confrontation articulée, où les modifications sont proposées, argumentées et discutées ensemble.

Le cadre théorique particulier se développe dans ces dialogues, mais il s'inscrit dans une formation psychologique et philosophique que j'ai acquise par moi-même bien avant de recourir à ces instruments.

La responsabilité des thèses soutenues, et de leurs faiblesses éventuelles, reste néanmoins la mienne.

Bruno Cancellieri

Préface

D'une personne qui rend toujours son travail en retard, on dit qu'elle est un procrastinateur. D'une entreprise qui ne respecte pas ses contrats, qu'elle est malhonnête. D'une mère qui téléphone trois fois par jour, qu'elle est envahissante. Ce sont des phrases qui semblent décrire et qui en réalité ferment : elles attribuent à quelqu'un une propriété stable, et d'une propriété stable il ne suit rien, sinon la possibilité de la constater de nouveau. Celui qui les prononce a cessé d'enquêter au moment même où il a cru avoir compris.

L'observation n'est pas de moi. Alfred Korzybski, dans *Science and Sanity* (1933), attribuait à un usage particulier du verbe être — celui qui identifie un individu à une catégorie, «X est un Y» — une part considérable de la confusion avec laquelle nous parlons du monde ; un de ses élèves, D. David Bourland, alla jusqu'à proposer d'écrire sans jamais employer ce verbe. Karl Popper, au paragraphe 10 de *Misère de l'historicisme*, appela essentialisme méthodologique l'attitude de qui commence par demander «qu'est-ce que la matière ?», «qu'est-ce que la justice ?», et nominalisme méthodologique celle de qui demande plutôt «comment se comporte ce morceau de matière ?» ; et il observa que la physique et la biologie ont fait leurs progrès en passant à la seconde question, tandis que les sciences de l'homme sont restées attachées à la première.

L'agentologie prend ce passage au sérieux et le porte là où il n'est pas encore tout à fait parvenu : au comportement des êtres vivants et de leurs organisations. À la question de ce qu'une chose est, elle substitue quatre questions préliminaires sur ce qu'elle fait : quels agents sont en jeu, ce que chacun perçoit, quel objectif il poursuit, quelles actions il tient pour disponibles. Elle a en cela un précédent illustre. Dans *La nature et la pensée* (1979), Gregory Bateson ne dit pas ce qu'est l'esprit : il énumère six conditions qu'un processus doit satisfaire pour qu'on puisse le dire mental, et la première est qu'il s'agisse d'un agrégat de parties en interaction, dont les interactions sont déclenchées par des différences et organisées en chaînes circulaires de détermination. Une définition entièrement faite de ce qui advient, et où le verbe être n'est pas employé.

L'avantage du passage de l'être à l'agir tient à la manière dont les deux descriptions finissent. Celle par identité ou par propriété finit en un verdict, et un verdict n'a pas de suite : il nomme quelque chose que l'on suppose constant, et ce qui est constant n'offre pas de prise. La description par agents ne peut pas finir ainsi, car chacune de ses quatre questions a une réponse qui pourrait être différente de ce qu'elle est : on peut percevoir autre chose, vouloir autre chose, disposer d'actions que l'on n'avait pas vues. Là où le verdict livrait un caractère, l'analyse livre quatre points où quelque chose peut être changé ; et dans le cas où aucun des quatre n'est modifiable, elle le dit, ce qui est de toute façon plus que ce qu'un verdict concède.

Il en découle deux choses. La première est qu'une description par agents reste démentissable. «Il est agressif» se présente comme un fait au sujet d'une personne, et n'admet pas de réplique ; «il perçoit mes questions comme des attaques» est une hypothèse sur une perception, que l'on peut vérifier en le lui demandant et, au besoin, corriger. La seconde est que le même vocabulaire traverse des domaines qui restent d'ordinaire séparés : les mêmes mots peuvent décrire une cellule, une personne, une entreprise et un programme informatique, et ce que l'on a appris à voir dans l'une de ces choses, on peut essayer de le voir dans les autres.

Tout cela ne fait pas de l'agentologie la meilleure manière de décrire n'importe quoi. Pour les phénomènes où l'on ne parvient à identifier ni frontière, ni perception, ni répertoire d'actions — l'érosion d'un rocher, la chute d'un corps pesant — la description par agents n'a pas de sens et n'est pas applicable ; le chapitre *Principe général et méthode agentologique* le dit en détail. La thèse est plus restreinte : lorsque la question porte sur un comportement, c'est-à-dire sur les raisons pour lesquelles quelqu'un fait ce qu'il fait et sur ce qui pourrait lui faire faire autre chose, la description par agents est préférable à celle par essences, et elle l'est pour la raison qui vient d'être dite.

L'agentologie est un appareil conceptuel de mon invention qui se propose de décrire et d'analyser, au moyen du concept d'agent, les phénomènes qui impliquent des êtres vivants, leurs organisations et les machines dotées d'un programme. Au lecteur qui n'a pas d'intérêt pour les appareils conceptuels, elle propose trois choses.

La première : lorsqu'on n'arrive pas à faire ce qu'on voudrait ou ce qu'on a décidé de faire, il ne s'agit pas d'un défaut de volonté. Ce sont deux parties de la même personne qui poursuivent des objectifs différents, et cette structure indique où l'on peut intervenir — sur la situation qui active le conflit, sur le signal par lequel l'une des deux parties se fait entendre, ou sur le programme qui engendre ce signal. Au lieu de penser à un défaut de caractère, on peut considérer trois points sur lesquels réfléchir.

La deuxième : les émotions ne sont pas des états de conscience à subir, mais des messages, avec un expéditeur, un objet et une demande. À une demande on peut répondre au lieu d'obéir ; et d'un message on peut établir quelle part est intelligible — ce que l'on sent l'est presque toujours, le pourquoi presque jamais.

La troisième : quand un rapport se détériore, presque jamais les deux n'y reconnaissent les mêmes figures, et presque jamais chacun ne fait ce que l'autre croit qu'il fait ; et il existe une opération qui permet de découvrir ensemble ce que chacun croyait être en train de faire : la méta-interaction.

Le centre de cet essai est le répertoire des figures relationnelles. Les chapitres qui le précèdent servent à le construire — ils disent ce qu'est un agent, comment il perçoit, comment il décide, et d'où lui viennent ses objectifs ; celui qui le suit le met à l'épreuve sur un cas. Sans le répertoire l'appareil resterait général et d'un emploi difficile ; sans l'appareil le répertoire serait une liste de noms.

Il faut dire tout de suite, cependant, que l'appareil est général et le répertoire humain : une cellule ne ment pas, ne manipule pas et ne se sent pas victime d'une injustice. L'appareil sert néanmoins, et c'est même lui qui permet de reconnaître les mêmes figures là où l'on ne s'attend pas à les trouver — la dépendance entre deux entreprises, le commandement entre deux programmes, la sélection entre deux organismes.

Le lecteur trouvera dans cet appareil des points de contact avec des traditions déjà existantes — la théorie générale des systèmes, la cybernétique, la notion d'agent rationnel de l'intelligence artificielle — qui ont déjà décrit des organismes, des machines et des organisations en termes de perception, de décision et d'action. L'agentologie ne revendique pas un champ d'études nouveau. Par rapport à ces traditions, elle fait deux choses : elle rend explicite un appareil qu'elles partagent souvent de manière seulement implicite et fragmentaire, et elle permet de l'appliquer à des domaines qui, à première vue, semblent n'avoir rien en commun.

Concepts fondamentaux

Les concepts qui suivent constituent le noyau théorique de l'agentologie. Agent, conscience, situation, perception, programme et action y sont assumés comme termes primitifs.

Agent

Un agent est une entité capable de percevoir la situation dans laquelle il se trouve et d'accomplir des actions qui tendent à la modifier ou à la conserver. Qu'il possède un état interne et au moins un programme ne fait pas partie de cette formulation : cela est fixé par les axiomes.

Un composant interne mérite le nom de sous-agent, et non simplement de programme, lorsqu'il manifeste avec une certaine persistance une ligne de conduite propre, non réductible à celle des autres composants, et qu'il la maintient lorsque les situations varient. Le critère est graduel et n'est pas toujours décidable : entre un programme très stable et un sous-agent peu autonome la différence est de degré, et dans bien des cas il n'y a pas moyen d'établir de quel côté tombe un cas particulier.

Un agent peut être constitué d'agents d'ordre inférieur, que l'on peut appeler sous-agents par rapport au premier. Chaque sous-agent offre des services à son agent supérieur sous forme d'informations (avertissements, demandes d'intervention et réponses à des questions). Un sous-agent peut en outre adresser des messages à d'autres sous-agents du même agent.

Conscience

La conscience est ce par quoi certaines informations, au lieu de se borner à produire des effets, sont senties. L'essai ne dit pas ce qu'elle est ni comment elle se produit : il l'assume comme primitive, au même titre qu'agent et que situation, et n'en emploie qu'une seule propriété — que ce qui l'atteint est senti. Son usage est contraint par l'axiome A1, qui attribue à tout être humain un seul agent doté de conscience et un nombre indéterminé d'agents qui en sont dépourvus.

Un agent est dit conscient s'il est doté de conscience, inconscient s'il ne l'est pas. Dans cet essai les deux termes sont toujours employés en ce sens, et jamais au sens plus faible d'«être au courant de quelque chose» : un agent inconscient peut disposer d'informations très exactes sur sa propre situation, et agir en fonction d'elles avec une grande efficacité.

Rien de ce qui suit ne dépend du fait de savoir ce qu'est la conscience. Cela dépend seulement du fait que, chez un être humain, certaines informations sont senties et d'autres non : ce qui est un constat, non une théorie. C'est sur ce seul fait que repose la définition de l'émotion donnée au chapitre *Les émotions*.

Situation

Une situation est la configuration, à un instant déterminé, de l'environnement extérieur et des états internes des agents impliqués. Puisque l'état interne comprend des informations, des mémoires et des programmes, entrent aussi dans la situation les relations entre les agents, qui n'ont pas d'existence propre mais consistent en ce que les interactions passées ont déposé dans les agents eux-mêmes. N'y entrent pas en revanche les interactions : une situation est instantanée, tandis qu'une interaction est une séquence d'actions, et c'est elle qui conduit d'une situation à la suivante.

Perception

La perception est le processus par lequel un agent acquiert des informations sur la situation et des messages provenant d'autres agents. Les perceptions modifient l'état interne de l'agent.

Programme

Un programme est une structure logique interne, partiellement stable dans le temps, qui contraint l'ensemble des actions qu'un agent peut produire à partir d'une perception donnée : elle exclut certaines actions et en rend d'autres plus probables ou nécessaires.

Cette définition a une conséquence précise : une entité dont les actions seraient entièrement indépendantes de toute structure interne — purement aléatoires par rapport à ce qu'elle perçoit — ne posséderait pas de programme et, par l'axiome A4 (voir la section suivante), ne serait pas un agent.

Constituent des exemples typiques de programmes les mécanismes génétiques, les réflexes, les instincts, les mécanismes qui engendrent les émotions, les habitudes, les connaissances, les stratégies, les raisonnements, les algorithmes, les logiques, etc.

Action

Une action est tout geste, mouvement ou expression — ou toute suite de ceux-ci — de la part d'un agent, visant à modifier une situation ou à la conserver. Toute action produit un ou plusieurs effets, intentionnels ou non intentionnels : la distinction entre les deux types d'effet est reprise dans la définition de Responsabilité, au chapitre *Définitions*.

Les actions ne restent pas isolées. Lorsque les actions de deux agents se répondent mutuellement, en modifiant réciproquement leurs états internes, on a une interaction. Lorsque les interactions entre les mêmes agents se répètent dans le temps sous des formes reconnaissables, ce qu'elles déposent dans leurs états internes est une relation. Les trois termes se situent sur une échelle de durée : l'acte singulier, l'échange où l'acte s'inscrit, le sédiment que les échanges répétés laissent.

Puisqu'une relation est un dépôt dans les états internes, et que les états internes font partie de la situation, les relations construites par les actions et les interactions passées font partie de la situation où s'accomplissent les actions suivantes.

Axiomes

Les énoncés qui suivent sont tout ce que l'essai assume. Aucun d'eux ne dit ce qu'un agent veut, ni pourquoi il poursuit un objectif plutôt qu'un autre : de telles questions relèvent du chapitre *Les émotions*.

A1. Il existe des agents : par exemple les organismes vivants, leurs organes, les cellules, les organisations et les machines dotées d'un programme. Tout être humain comporte un agent conscient — sa conscience — et un nombre indéterminé d'agents inconscients.

Remarque. Que tout être humain comporte un seul agent conscient et un nombre indéterminé d'agents inconscients est une thèse sur l'être humain, non falsifiable en l'état actuel, dont j'assume la responsabilité. Je la place parmi les axiomes parce que le chapitre *Les émotions* repose entièrement sur elle.

A2. Tout agent se trouve toujours dans une situation.

A3. Tout agent perçoit au moins une partie de la situation dans laquelle il se trouve.

A4. Tout agent possède au moins un programme.

A5. Le comportement d'un agent à un instant donné dépend de la situation perçue et de son état interne.

A6. Les programmes d'un agent peuvent se modifier dans le temps (programmes plastiques) ou rester invariants (programmes rigides).

A7. Un agent peut, par ses propres actions, modifier la situation, d'autres agents ou son état interne, et en particulier ses programmes.

A8. Un ensemble d'agents constitue un agent d'ordre supérieur lorsque, au niveau de l'ensemble, émergent des perceptions et des décisions non réductibles à la simple somme des perceptions et des décisions des composants pris un à un — par exemple une procédure de décision coordonnée, une frontière reconnaissable par rapport à l'extérieur, une persistance dans le temps distincte de celle des composants. En ce cas les agents peuvent être considérés comme des sous-agents par rapport à l'agent d'ordre supérieur. En l'absence de ces conditions, l'ensemble reste une pluralité d'agents, et non un agent.

A9. Un agent peut appartenir simultanément à plusieurs agents d'ordre supérieur.

Définitions

Information — Toute différence relevée et acquise au moyen d'une perception et enregistrée dans l'état interne d'un agent, qui peut donner lieu à des différences de comportement ou qui peut être émise dans l'environnement vers d'autres agents.

État interne — Ensemble des informations, des programmes et des autres caractéristiques internes possédées par un agent à un instant déterminé.

Décision — Opération par laquelle un programme sélectionne une action parmi celles que la situation perçue rend disponibles (cf. Liberté). Le programme est la structure, la décision est l'opération que cette structure accomplit : c'est pourquoi ils figurent comme deux moments distincts dans le cycle fondamental. Les décisions sont prises selon des logiques qui, sous leur forme la plus élémentaire, consistent en l'instruction «si X, alors Y», où X est une situation reconnue et Y l'action à exécuter.

Objectif — Situation qu'un agent tend à réaliser, à maintenir ou à éviter.

Mémoire — Partie de l'état interne qui conserve des informations dans le temps.

Apprentissage — Modification d'un ou de plusieurs programmes à la suite de l'expérience ; c'est le cas, prévu par A6, d'un programme plastique qui se modifie. Ce n'est pas un phénomène exclusivement humain ni exclusivement psychologique : un lymphocyte mémoire modifie son propre programme à la suite d'une exposition et conserve la modification pendant des années.

Interaction — Séquence d'actions par laquelle deux agents ou plus modifient réciproquement leurs états internes respectifs.

Communication — Interaction dans laquelle au moins une action a pour objectif spécifique, et non comme simple effet collatéral, la modification de l'état interne d'un autre agent au moyen d'un signal régi par un code au moins en partie partagé entre les agents impliqués. Cette clause distingue la communication de l'ensemble des interactions : toute interaction qui modifie l'état interne d'autrui n'est pas une communication, mais seulement celle où cette modification est poursuivie comme objectif au moyen d'un code.

Commandement — Communication dont l'objectif est que le destinataire exécute une action visant à obtenir une certaine situation.

Coopération — Interaction dans laquelle deux agents ou plus poursuivent au moins un objectif commun et où leurs actions respectives sont réciproquement fonctionnelles à son obtention.

Compétition — Interaction dans laquelle l'obtention de l'objectif d'un agent réduit la probabilité qu'un autre agent obtienne le sien.

Remarque. Coopération et compétition ne sont pas des catégories mutuellement exclusives : une même interaction peut être coopérative quant à un objectif et compétitive quant à un autre (interaction mixte) — c'est le cas ordinaire des relations entre organisations qui collaborent sur un projet et se font concurrence sur un marché.

Pouvoir — Capacité d'un agent à modifier de manière significative et durable l'état ou le comportement d'un autre agent, même contre les préférences de ce dernier ou en surmontant la résistance qu'il oppose éventuellement. Cette clause distingue le pouvoir de la capacité d'action en général : toute action qui modifie un autre agent n'est pas un exercice de pouvoir, mais seulement celle qui l'emporte sur une préférence ou une résistance contraire.

Liberté — Ensemble des actions qu'un agent tient pour disponibles dans la situation où il se trouve et qu'il est en mesure d'exécuter avec les programmes dont il dispose. La liberté ainsi définie est relative au jugement de l'agent : une action objectivement possible mais qu'il ne tient pas pour telle n'en est pas, pour cet agent, dans sa liberté. La divergence entre les actions tenues pour disponibles et celles qui sont objectivement disponibles pour un observateur extérieur est elle-même un phénomène agentologiquement pertinent — par exemple dans les cas d'ignorance ou d'auto-illusion.

Responsabilité — Attribution à un agent des effets de sa propre action, graduée selon deux facteurs : le degré d'intentionnalité de l'effet (cf. Action) et l'étendue des alternatives concrètement disponibles pour lui au moment de la décision. Un effet pleinement intentionnel, produit en présence d'alternatives disponibles, fonde une attribution de responsabilité pleine ; l'absence d'alternatives ou d'intentionnalité l'atténue ou l'exclut. L'attribution peut avoir pour objet un autre agent ou l'agent lui-même. C'est l'opération sur laquelle se fondent les jugements moraux et les évaluations pratiques : on réagit différemment au même dommage selon qu'on le juge intentionnel ou accidentel.

Remarque. Ici « disponibles » n'a pas le même sens que dans la définition de Liberté : là il s'agissait des actions que l'agent tenait pour telles, ici de celles qu'un observateur reconnaît comme possibles pour lui. Que l'agent ne les ait pas vues peut atténuer l'attribution, mais ne suffit pas à l'exclure.

Situation relationnelle — Partie de la situation qui concerne deux agents déterminés : la configuration de leurs objectifs, de leurs programmes réciproquement pertinents, de l'asymétrie entre eux et des effets que l'un produit sur l'autre, à un instant donné. Elle est instantanée comme toute situation, et elle subsiste aussi entre des agents qui se rencontrent pour la première fois ; ce que les situations relationnelles successives déposent dans les états internes est la relation.

Relation — Lien relativement stable entre deux agents ou plus, constitué par la sédimentation dans le temps d'une multiplicité d'interactions entre eux, de manière à orienter de façon reconnaissable les interactions suivantes. Une relation est constituée par les habitudes interactives — partagées ou réciproquement représentées — déposées dans les états internes des agents impliqués. En ce sens, une relation pourrait être considérée comme le « pattern » (c'est-à-dire l'ensemble des formes caractéristiques et répétitives) des interactions habituelles entre deux agents. Les formes qui y récurrent sont des figures relationnelles, définies à l'entrée suivante. Elle doit être distinguée de la situation relationnelle, dont elle est le sédiment.

Figure relationnelle — Configuration reconnaissable que deux agents réalisent ensemble, sous un aspect déterminé : les objectifs des deux, leurs programmes, l'asymétrie entre eux, la valence affective, les effets que l'un produit sur l'autre. Coopération, dépendance, commandement et obéissance sont des figures relationnelles. Comme les aspects sont nombreux, une même situation relationnelle en présente simultanément plusieurs — deux collègues peuvent coopérer, dépendre l'un de l'autre et ne pas se faire confiance ; et comme chacun la reconnaît pour son propre compte, une même figure donne lieu à deux lectures, une par agent, qui peuvent diverger. Ce sont les figures reconnues, et non la situation relationnelle, qui entrent dans les décisions.

Remarque. Une figure se reconnaît dans une action isolée, dans une interaction ou dans une relation ; les figures qui récurrent dans une relation sont les formes caractéristiques dont son pattern est fait. Dans cet essai on écrit figure relationnelle dans les définitions, dans les titres et à la première occurrence de chaque section, et figure là où le contexte est clair.

Conflit — Situation dans laquelle deux agents ou plus, ou deux programmes ou plus d'un même agent, tendent vers des actions ou des objectifs mutuellement incompatibles ou antagonistes. La compétition est le cas où les termes du conflit sont des agents distincts ; le conflit intra-agent est le cas où ce sont des composants d'un même agent.

Principe général et méthode agentologique

Tout phénomène qui présente une configuration non aléatoire de perception, de décision et d'action peut être décrit comme le résultat des actions d'un ou de plusieurs agents qui, sur la base de leurs propres perceptions et de leur état interne — lequel comprend programmes et mémoire — modifient ou maintiennent, selon certaines logiques, les situations dans lesquelles ils sont immergés.

Ce principe ne s'applique pas aux processus dépourvus de rétroaction, où rien de ce qui advient ne dépend de ce que quelqu'un perçoit — par exemple l'érosion d'un rocher sous l'effet du vent. En ces cas l'agentologie n'a rien à ajouter à la description physique ordinaire, et ne prétend pas s'y substituer.

La méthode agentologique consiste à répondre aux questions contenues dans le questionnaire suivant et à d'autres, plus détaillées, qui peuvent en être dérivées.

Questionnaire agentologique

- Quels sont les agents impliqués ?
- Quelle est la situation actuelle ?
- Quelles informations chaque agent perçoit-il ?
- Quelles figures chaque agent reconnaît-il entre lui et chacun des autres ?
- Quels programmes sont pertinents ?
- Quels objectifs chaque agent poursuit-il ?
- Quelles actions accomplit-il ?
- Selon quels critères et selon quelles logiques chaque agent décide-t-il de ses propres actions ?
- De quelles manières ces actions modifient-elles la situation ?
- Quelles nouvelles situations peuvent en être engendrées ?

Les questions suivent l'ordre du cycle fondamental de l'agentologie :

Situation → Perception → Programme → Décision → Action → Nouvelle situation

La dernière question ramène à la première : la situation produite par les actions est celle à partir de laquelle l'analyse recommence. La méthode n'est donc pas une liste de réponses, mais un tour à parcourir autant de fois qu'il y a de situations qui se succèdent.

Les émotions

Le noyau axiomatique défini plus haut attribue à tout agent des objectifs, mais ne dit pas par quoi un objectif est déterminé, ni pourquoi un agent en poursuit un plutôt qu'un autre. L'axiome A1, en outre, affirme que tout être humain comporte un agent conscient — sa conscience — et un nombre indéterminé d'agents inconscients, mais rien n'a encore été dit sur la manière dont ceux-ci communiquent entre eux. Les deux questions ont, à mon avis, la même réponse, et c'est pourquoi ce chapitre les aborde ensemble : les objectifs que la conscience poursuit lui sont en grande partie proposés, ou imposés, par les agents inconscients, et le moyen dont ceux-ci disposent pour le faire, ce sont les émotions. Les émotions sont, en ce sens, le principal mécanisme par lequel un être humain a les objectifs qu'il a et non d'autres.

Le sens des émotions

J'appelle ici émotion seulement ce qui est senti. Un processus affectif qui n'atteint pas la conscience — une activation physiologique, une évaluation automatique, une préférence qui oriente le comportement sans que le sujet s'en aperçoive — n'est pas, dans ce vocabulaire, une émotion : c'est un programme, ou l'opérer d'un agent inconscient. Il s'agit d'un choix terminologique, non d'une thèse empirique. Il ne nie pas l'existence des processus affectifs qui n'atteignent pas la conscience, lesquels sont au contraire le présumé de tout le chapitre ; il réserve seulement le nom d'émotion au signal qui, de ces processus, parvient à la conscience.

L'émotion comme message

Émotion — Information qu'un agent inconscient transmet à la conscience, et qui est sentie par celle-ci en termes de plaisir ou de douleur, d'attraction ou de répulsion. Elle est dotée de trois attributs : un objet (une situation ou un événement présent, remémoré ou imaginé), une valence positive ou négative envers cet objet, et une intensité.

Une émotion est donc une communication, c'est-à-dire une action dont l'objectif spécifique, et non le simple effet collatéral, est de modifier l'état interne de l'agent «conscience». C'est cependant une communication dont le code n'est que partiellement partagé. Valence et intensité parviennent à la conscience avec une relative clarté : on sent qu'on est attiré ou repoussé, et on sent à quel point. L'objet et la raison peuvent parvenir déformés, mal reconnus, ou ne pas parvenir du tout. C'est l'expérience ordinaire de sentir quelque chose sans savoir pourquoi, et c'est la conséquence prévisible du fait que l'expéditeur n'est pas conscient et n'est pas en mesure de construire et d'envoyer à la conscience des messages plus clairs et plus éloquents.

Reconstruire le message complet — quel agent l'émet, sur quel objet, en vue de quel objectif — est une opération d'abduction, la même qui, au chapitre *Application à la psychothérapie et à l'autothérapie*, est assimilée au reverse engineering. Sa faillibilité n'est pas un accident, mais une caractéristique structurelle, et elle doit être présente à l'esprit chaque fois que quelqu'un — le sujet lui-même ou un observateur — déclare savoir pourquoi une personne éprouve ce qu'elle éprouve.

La force du message : demandes et commandements

Une émotion n'est pas une information neutre. Elle est accompagnée d'une promesse : de plaisir, si la conscience suit l'attraction ou évite ce qui la repousse ; de douleur, dans le cas contraire. C'est la promesse qui donne au message sa force, et qui fait d'une information une sollicitation à agir.

Selon son intensité, une émotion fonctionne comme demande ou comme commandement impératif, au sens où le commandement a été défini au chapitre *Définitions*. À faible intensité elle laisse à la conscience le temps et l'espace de délibérer, de différer, de décider autrement : elle entre dans la délibération comme l'un de ses éléments, et en ce sens elle est une demande. À forte intensité la

délibération est exclue, ou bien elle a lieu mais son issue est sans importance parce que l'action est déjà en cours : c'est un commandement, et c'est le cas où l'agent inconscient exerce un pouvoir sur la conscience. Entre les deux extrêmes il n'y a pas de frontière mais un continuum.

Cette distinction fournit la description qui, autrement, fait défaut à l'acrasie — le manque de maîtrise de soi. La décrire, comme on le fera au chapitre *Application à la psychothérapie et à l'autothérapie*, comme la prévalence d'un programme sur un autre est correct, mais incomplet : les programmes, pris en eux-mêmes, n'ont ni préférences à faire valoir ni moyens de les faire valoir. Ce qui prévaut, c'est un programme qui dispose d'un canal pour atteindre la conscience et d'une promesse pour faire pression sur elle.

L'intensité de l'émotion n'a pas toujours le même effet. La même intensité est résistée différemment par des personnes différentes, et par la même personne à des moments différents — selon la fatigue, l'accoutumance à ce signal particulier, la présence d'un signal contraire. L'intensité est un facteur nécessaire pour rendre compte de la force d'une émotion, non suffisant : ce qu'une intensité donnée parvient à obtenir dépend aussi de l'état interne de celui qui la reçoit.

Attraction et répulsion

Toute émotion comporte une attraction ou une répulsion envers son objet. Dans le vocabulaire de cet essai : elle propose à la conscience un objectif — réaliser ou maintenir la situation dont l'objet fait partie, dans le cas de l'attraction ; l'éviter ou l'interrompre, dans le cas de la répulsion.

Les émotions mixtes — celles où le même objet attire et repousse à la fois — ne font pas exception à cette thèse : elles en sont une conséquence. Ce sont deux messages, émis par deux agents inconscients différents, qui atteignent simultanément la conscience avec des valences opposées ; c'est-à-dire un cas de conflit intra-agent, tel qu'il est décrit au chapitre suivant. Elles ont toutefois une particularité qui les rend précieuses : ce sont le seul type de conflit interne dont les deux termes se présentent l'un et l'autre à la conscience. Les autres conflits intra-agents doivent être reconstruits par abduction ; l'ambivalence, elle, se sent. C'est probablement la raison pour laquelle il convient que l'auto-analyse commence par là.

Restent des cas que cette description ne couvre pas. La surprise ne semble pas avoir de valence propre : elle précède l'évaluation au lieu de l'exprimer, et ressemble davantage à un avertissement — l'information par laquelle un sous-agent signale que quelque chose a changé, sans pour autant demander une action déterminée — qu'à une émotion au sens ici défini. Les états d'âme diffus, qui semblent dépourvus d'objet, admettent deux lectures : ou bien l'objet existe mais il est si général qu'il n'est pas identifiable, ou bien ce ne sont pas des émotions mais le bruit de fond produit par l'état interne d'ensemble. Je laisse la question ouverte.

Ce que la conscience peut faire du message

Si le rapport était à sens unique — des commandements qui arrivent, une exécution qui suit — il n'y aurait rien à dire du changement, et le chapitre *Application à la psychothérapie et à l'autothérapie* resterait sans objet. Les actions dont la conscience dispose sont au moins celles-ci : exécuter ; différer l'exécution, ce qui est déjà une forme de résistance ; agir non pas sur l'action demandée mais sur la situation qui engendre le signal, en se soustrayant aux contextes qui l'activent ; interroger le signal, c'est-à-dire tenter la reconstruction du message complet ; et enfin modifier le programme qui l'émet, ce qui est la voie la plus lente et la seule durable.

Chacune de ces actions est à son tour perçue par les agents inconscients, qui recalibrent en conséquence leurs propres signaux. Le rapport est donc un cycle, non une chaîne de commandement : c'est la même structure Situation → Perception → Programme → Décision → Action → Nouvelle situation, appliquée à l'intérieur d'une seule personne.

Une conséquence mérite d'être signalée parce qu'elle touche un point du chapitre *Concepts fondamentaux*, où il a été dit que chaque sous-agent offre des services à son agent supérieur sous forme d'informations. Si les agents inconscients commandent la conscience, la conscience n'est pas leur agent supérieur, mais leur pair : l'agent supérieur est la personne dans son ensemble, dont tant la conscience que les agents inconscients sont des sous-agents, et les émotions sont des communications horizontales entre sous-agents, non des services rendus vers le haut. La conscience n'est pas le sommet de la personne : elle est son sous-agent spécialisé dans la délibération et dans l'action à longue portée, et elle est à ce titre destinataire des messages des autres.

Ce qui change, à les traiter ainsi

La différence entre un état et un message n'est pas verbale. D'un état on ne peut qu'attendre qu'il change, ou tenter de le supprimer : ce sont les deux choses que l'on fait communément avec la peur, avec la colère, avec le désir inopportun. À un message, en revanche, on peut répondre, et les réponses disponibles sont celles qui ont été énumérées à la section précédente. Aucune d'elles n'est disponible pour qui considère sa propre émotion comme une condition dans laquelle il se trouve.

On en tire deux choses. La première est que l'on peut distinguer la part fiable du signal de celle qui ne l'est pas. Ce que l'on sent — attraction ou répulsion, et avec quelle force — arrive presque toujours correct : c'est l'information que l'expéditeur sait donner. Ce à quoi on l'attribue arrive souvent faux. «Je suis anxieux» est presque toujours vrai ; «je suis anxieux à cause de la réunion de demain» est une hypothèse, et se demander si c'est la bonne est une opération qui n'a de sens que pour qui a admis que l'émotion a un objet, et que cet objet pourrait être un autre.

La seconde est que l'ambivalence cesse d'être un défaut. Qui traite les émotions comme des états considère avec embarras le fait de ne pas savoir s'il veut ou non une chose, et cherche à le résoudre en décidant, c'est-à-dire en faisant taire l'un des deux côtés. Qui les traite comme des messages y lit une information : deux expéditeurs différents, avec des demandes opposées, dont chacun protège quelque chose. La question utile n'est plus de savoir lequel des deux a raison, mais ce que chacun cherche à protéger — et c'est une question à laquelle, contrairement à l'autre, il arrive qu'on sache répondre.

Les figures relationnelles

Le chapitre *Définitions* a distingué trois choses que l'usage courant confond : la situation relationnelle, qui est la configuration instantanée entre deux agents ; la relation, qui est ce que les situations relationnelles successives déposent dans leurs états internes ; et la figure relationnelle, qui est une configuration reconnaissable que les deux réalisent ensemble sous un aspect déterminé. Ce chapitre énumère les figures qui me semblent les plus pertinentes pour les décisions, et montre ce qui découle du fait que chacun des deux agents les reconnaît pour son propre compte.

La situation relationnelle et les figures

D'un même couple d'agents on peut donner des descriptions différentes. L'une est celle d'un tiers observateur, qui compare les objectifs des deux et constate, par exemple, que l'obtention de l'un réduit la probabilité que l'autre obtienne le sien : c'est la compétition telle que le chapitre *Définitions* la définit, et c'est une configuration du couple, non une propriété de l'un des deux. Les autres sont les figures relationnelles que les deux agents y reconnaissent : la manière dont chacun d'eux lit cette configuration. Les trois descriptions peuvent ne pas coïncider, et en règle générale elles ne coïncident pas entièrement ; et la première non plus n'est pas une description faite de nulle part, car le tiers observateur est lui aussi un agent qui perçoit, mieux informé mais non moins situé.

Il ne s'agit pas d'une difficulté propre aux figures. C'est l'application à la situation relationnelle de ce que l'essai a déjà établi pour la situation et pour la liberté. L'axiome A5 dit que le comportement d'un agent dépend de la situation perçue, et la définition de la perception explique pourquoi : à la situation aucun agent n'a accès, il a accès aux informations que sa perception lui livre. La situation relationnelle est une partie de la situation, et elle ne fait pas exception : elle n'entre dans le comportement qu'à travers les figures qu'on y reconnaît. La divergence entre la situation relationnelle et la figure reconnue est la même que celle qui, dans la définition de la liberté, sépare les actions objectivement disponibles de celles que l'agent tient pour telles.

De là vient une conséquence qu'il vaut la peine d'énoncer à part : les deux lectures d'une même figure peuvent être asymétriques, et leur asymétrie n'est pas un cas pathologique mais la condition ordinaire. L'un croit rendre un service, l'autre reçoit un commandement. L'un coopère, l'autre se sent exploité. Aucun des deux ne ment.

La figure relationnelle comme entrée de la décision

Reconnaître une figure relationnelle n'est pas une conclusion, c'est une entrée. Dans le cycle fondamental — Situation → Perception → Programme → Décision → Action → Nouvelle situation — la reconnaissance de la figure n'ajoute pas une phase : elle fait partie de ce que la perception livre aux programmes. Les règles de la forme «si X, alors Y» que le chapitre *Définitions* attribue à la décision ont souvent pour X non pas une situation physique mais une figure : si celui-ci est un concurrent, alors je ne lui dis pas ce que je sais ; s'il a autorité sur moi, alors j'exécute ; si je dépends de lui, alors je ne proteste pas.

Il s'ensuit une propriété qui distingue les figures des autres perceptions, et c'est la raison principale pour laquelle elles méritent un chapitre. La figure reconnue tend à se rendre vraie. Celui qui reconnaît qu'il est en compétition agit en conséquence, l'autre perçoit ces actions et répond à son tour de façon compétitive, et la situation relationnelle finit par ressembler à la figure qui l'a précédée. C'est le même cycle qu'auparavant, parcouru une fois de plus ; mais son effet est qu'une erreur de reconnaissance ne reste pas une erreur de description : elle devient une action qui modifie ce qu'elle décrivait. Au chapitre *Un exemple d'application*, c'est ce qui advient entre Marco et Giulia : l'intensification des contrôles, décidée à partir d'une figure où Giulia lit un désengagement, produit le désengagement qu'elle présumait.

C'est aussi la réponse à l'objection que ce chapitre s'attire. Dire «notre rapport est un rapport de compétition» semble être le même geste que la *Préface* déclare stérile : attribuer une propriété stable et clore l'enquête. La différence tient à ce que l'on suppose avoir reconnu. Une figure n'est pas une propriété d'une personne : c'est une hypothèse sur une configuration entre deux, que l'on peut vérifier en demandant et, au besoin, corriger ; et à la différence d'un caractère elle a une date de début, une histoire, et au moins deux lectures. Un verdict est stérile lorsqu'il se croit unique et définitif, non du fait d'être prononcé.

La reconnaissance des figures relationnelles

Percevoir une figure relationnelle et la reconnaître ne sont pas la même opération. La perception, telle que le chapitre *Concepts fondamentaux* la définit, est le processus par lequel un agent acquiert des informations sur la situation ; la reconnaissance est la confrontation de ces informations avec un répertoire que l'agent possède déjà. Ce répertoire est fait de programmes, au sens de l'entrée correspondante, qui énumère parmi ses exemples les connaissances, les stratégies et les logiques ; et puisque l'axiome A6 distingue les programmes rigides des programmes plastiques, certaines figures se reconnaissent sans les avoir jamais apprises et d'autres seulement pour les avoir rencontrées.

Une figure se reconnaît à trois niveaux, et non avec la même fiabilité : dans une action isolée — un ton, une correction, un silence ; dans une interaction, c'est-à-dire dans un échange qui se déroule ; et dans une relation, où ce que l'on reconnaît est ce qui récurse. Attribuer une figure à une relation sur la preuve d'un seul acte est la manière ordinaire de se tromper, et c'est le même saut que la *Préface* reproche à qui, de quelques remises en retard, conclut qu'une personne est un procrastinateur. L'échelle à laquelle on reconnaît est donc, à côté de celles déjà nommées, une source d'erreur.

Il s'ensuit une seconde raison pour laquelle deux agents peuvent diverger, indépendante de la première. On a dit jusqu'ici que chacun a accès à ses propres intentions et aux effets qu'il subit, et non à ceux de l'autre. Mais les deux peuvent aussi avoir des répertoires différents : une figure que l'un reconnaît peut manquer entièrement à l'autre, qui ne la lit pas mal, il ne la lit pas du tout. C'est le cas où une méta-interaction ne suffit pas, car il ne s'agit pas de corriger une attribution mais de montrer une figure que l'autre n'a jamais appris à voir.

Enfin, les figures ne se reconnaissent pas seulement : elles se poursuivent. Puisqu'une figure fait partie d'une situation, elle peut être un objectif au sens de l'entrée correspondante — une situation qu'un agent tend à réaliser, à maintenir ou à éviter. C'est ce qui rend mécanique le cycle énoncé à la section précédente : celui qui reconnaît une figure de compétition se donne pour objectif de la réaliser ou de la gagner, agit en conséquence, et l'autre la reconnaît à son tour. Les figures sont ainsi à la fois une entrée et une sortie du processus décisionnel : elles se reconnaissent dans la situation et se poursuivent dans l'action.

Répertoire des figures relationnelles

Les figures relationnelles qui suivent sont au nombre de vingt-deux. Elles ne sont pas une classification, et ne prétendent rien épuiser : elles sont un répertoire, c'est-à-dire la liste de ce qu'entre deux agents il vaut la peine de savoir reconnaître. Le chapitre précédent a dit ce qu'est une figure et comment on la reconnaît ; celui-ci dit lesquelles.

Ce qu'il faut attendre de la liste

Les entrées qui suivent ne s'excluent pas mutuellement. Un couple d'agents reconnaît presque toujours plusieurs de ces figures relationnelles à la fois, et d'autres successivement : deux collègues peuvent coopérer sur un projet, se disputer une promotion, dépendre l'un de l'autre de façon asymétrique et ne pas se faire confiance, le tout dans le même mois. Certaines entrées sont indépendantes les unes des autres, d'autres non, et là où le lien existe il est signalé : la dépendance est le fondement du pouvoir, le partage est une condition de la coopération, l'attraction oriente la sélection sans la déterminer.

La liste est ouverte. Ce n'est pas une classification exhaustive et elle ne dérive pas d'un principe qui en garantirait la complétude : elle rassemble les figures qui me semblent le plus fréquemment pertinentes pour les décisions, et d'autres peuvent y être ajoutées pourvu qu'elles passent le critère énoncé au chapitre *Un exemple d'application* — qu'une distinction vaut ce qu'elle coûte lorsque quelque chose change, soit l'action que l'on choisit, soit ce que l'on s'attend à voir advenir.

Enfin, toutes les figures ne sont pas applicables à tous les types d'agents. Chacune exige de l'agent qui la reconnaît certaines capacités, et là où celles-ci manquent la figure n'est pas rare : elle est dépourvue de sens. Attribuer une offense exige de savoir attribuer des intentions ; faire confiance exige de se représenter le comportement futur d'un autre ; imiter exige des programmes plastiques. Une bactérie se fait concurrence et coopère, elle ne s'offense pas ; un thermostat ne fait ni l'un ni l'autre. Là où la restriction compte, les entrées l'indiquent.

Sommaire des figures relationnelles

- Coopération, compétition, indifférence
- Hostilité
- Partage, non-partage
- Engagement
- Imitation, différenciation
- Connaissance de l'autre
- Sincérité, mensonge
- Vérité, fausseté
- Manipulation
- Confrontation et évitement
- Dépendance, indépendance
- Pouvoir exercé, pouvoir subi
- Commandement, obéissance
- Supériorité et infériorité reconnues
- Supériorité et infériorité institutionnelles
- Donner et recevoir un service
- Dette et gratitude
- Attraction, répulsion
- Confiance, méfiance
- Dommage, offense
- Justice, injustice
- Choisir et écarter, être choisi et être écarté

Dans les sections qui suivent les figures sont regroupées par catégorie et définies.

Les figures relationnelles quant aux objectifs

Coopération, compétition, indifférence — Les deux premiers sont définis au chapitre *Définitions*, avec l'observation qu'ils ne s'excluent pas : une même interaction peut être coopérative quant à un objectif et compétitive quant à un autre. Il faut ajouter ici le troisième cas, que la liste sans lui laisserait à découvert : l'indifférence, où les objectifs des deux n'interfèrent pas du tout. C'est la figure la plus commune entre des agents quelconques, et c'est celle que l'on perd de vue en premier, parce que celui qui se sent en compétition tend à attribuer à l'autre un intérêt pour lui que l'autre n'a pas. La décision qui change est de savoir s'il vaut la peine d'agir sur l'autre : d'un agent indifférent il n'y a rien à obtenir, et croire le contraire est l'une des erreurs les plus coûteuses.

Hostilité — Figure dans laquelle le fait de nuire à l'autre n'est pas un effet des actions mais leur objectif. Elle se distingue de la compétition par la structure de l'objectif et non par l'intensité : dans la compétition la perte de l'autre est la conséquence de son propre gain et cesse une fois le gain obtenu ; dans l'hostilité c'est la diminution de l'autre qui est poursuivie, et elle peut continuer quand on n'en tire plus rien. La guerre en est le cas institutionnel et organisé, mais la figure se rencontre entre deux personnes et entre deux organisations qui ne se combattent pas avec des armes. Elle doit aussi être tenue distincte de *Dommage, offense*, qui concerne un effet déjà produit et son attribution : l'hostilité est un objectif, et peut exister avant qu'aucun dommage n'ait été fait. Le désalignement typique est entre celui qui se sent en compétition et celui qui se sent attaqué, et par la propriété énoncée plus haut il tend à se résoudre dans le second sens : qui est traité en ennemi se comporte en ennemi. La décision qui change est de savoir si le conflit finira quand l'objet disputé sera attribué — dans la compétition oui, dans l'hostilité non.

Partage, non-partage — Deux agents peuvent partager des informations, des ressources, des critères de jugement, des valeurs, ou l'appartenance à un tiers agent. L'appartenance doit toutefois être distinguée selon son objet : de deux membres d'une même organisation on attend un certain comportement parce qu'une procédure de décision commune les lie ; de deux membres d'une même catégorie — qui n'est pas un agent et ne décide rien — on attend quelque chose seulement parce qu'on les a rangés dans une classe, et cette seconde attente a presque toujours un fondement plus faible. Le partage est une condition de la coopération — on ne coopère pas sur un objectif que l'un des deux ignore, ni avec des critères de succès incompatibles — mais il ne suffit pas à la produire. Et il peut produire le contraire : lorsque ce qui est partagé est le désir d'un objet qui n'admet pas deux possesseurs, le partage engendre la compétition, et d'autant plus âpre que les deux se ressemblent davantage. C'est la thèse que René Girard a développée dans *Mensonge romantique et vérité romanesque* (1961) : le désir ne naît pas de l'objet, mais de l'imitation d'un autre qui le désire, et le modèle devient par là même un rival. La décision qui change concerne ce qu'il faut mettre en commun : augmenter le partage améliore la coopération là où les objets sont divisibles, et la détériore là où ils ne le sont pas.

Engagement — Figure dans laquelle un agent lie son propre comportement futur à l'égard d'un autre : la promesse, le contrat, la parole donnée. Dans les termes de l'essai c'est un agent qui restreint volontairement sa propre liberté — l'ensemble des actions qu'il tient pour disponibles — en échange de la prévisibilité de l'autre ; et quand l'engagement est réciproque, tous deux achètent de la prévisibilité avec de la liberté. Il doit être tenu distinct de la confiance, qui est l'attente de celui qui reçoit et non l'acte de celui qui se lie : on s'engage aussi envers qui l'on ne croit pas. Le désalignement est quotidien et coûteux, car les engagements ont des degrés que les mots ne distinguent pas : l'un croit avoir dit «en principe», l'autre a déjà compté sur lui et renoncé à des alternatives. La décision qui change est de savoir s'il est encore licite d'y revenir, et à quelles conditions : devant une intention on change d'avis, devant un engagement on négocie ou l'on manque à sa parole.

Les figures relationnelles quant aux programmes et à leur connaissance

Imitation, différenciation — Un agent peut faire siens les programmes d'un autre, ou s'en éloigner délibérément. L'imitation exige des programmes plastiques, au sens de l'axiome A6, et présuppose un jugement : on imite celui que l'on juge accompli, et c'est pourquoi imiter est à la fois un acte technique et une déclaration de rang. La différenciation a deux mobiles qu'il convient de ne pas confondre : éviter le recouvrement, lorsque la ressemblance produit une compétition pour les mêmes ressources, et affirmer une identité, lorsque ce que l'on veut est de ne pas être confondu. Les deux côtés voient souvent la même chose de manière opposée : celui qui imite l'entend comme une reconnaissance, celui qui est imité peut y lire une appropriation. La décision qui change est de se rapprocher ou de s'éloigner des programmes de l'autre, et la réponse dépend duquel des deux mobiles est en jeu.

Connaissance de l'autre — C'est le degré auquel un agent dispose d'un modèle des programmes, des perceptions et des objectifs d'un autre. Elle n'est pas binaire et elle n'est pas symétrique : on peut très bien connaître un agent dont on est très peu connu, et c'est l'une des formes les plus stables d'avantage. Cette figure relationnelle n'exige aucune ressemblance ni aucune sympathie — on modélise un adversaire précisément pour ne pas faire ce qu'il ferait — et elle doit être tenue distincte de l'imitation, qui comporte d'adopter les programmes d'autrui et non simplement de se les représenter. C'est de modèles de ce genre que sont faits les sous-agents intériorisés dont parlent les chapitres *Application aux relations sociales* et *Application à la psychothérapie et à l'autothérapie*, et c'est sur eux qu'opère la méta-interaction. La décision qui change est de savoir dans quelle mesure se fier à ses propres prévisions sur le comportement de l'autre.

Les figures relationnelles quant à la communication

Sincérité, mensonge — Figure dans laquelle ce qu'un agent transmet correspond ou non à ce qu'il croit lui-même. Elle concerne l'expéditeur et non le contenu : un agent sincère peut transmettre le faux, si c'est lui qui est dans l'erreur ou qui se trompe lui-même — l'auto-illusion que le chapitre *Définitions* nomme à propos de la liberté et sur laquelle revient le chapitre *Application à la psychothérapie et à l'autothérapie*. Elle comprend aussi le silence : on ment en taisant ce que l'on sait pertinent pour l'autre, et l'omission en est la forme la plus commune et la moins reconnaissable. Elle doit être tenue distincte de la méfiance, qui est l'attente de celui qui reçoit et non l'acte de celui qui transmet : on peut se fier à un menteur et se méfier d'un sincère. C'est l'une des deux seules figures de la liste où la divergence entre les deux lectures n'est l'erreur d'aucun des deux, mais le résultat voulu par l'un. La décision qui change est de savoir s'il vaut mieux interroger l'autre ou vérifier par soi-même : à un menteur les questions ne servent pas.

Vérité, fausseté — Figure dans laquelle ce qui passe entre les deux correspond ou non à l'état des choses. Elle est indépendante de la précédente, et les quatre combinaisons se rencontrent toutes, y compris celle du menteur qui dit par hasard le vrai. C'est la seule figure de la liste dont les deux lectures peuvent coïncider et être toutes deux fausses : deux agents parfaitement sincères l'un envers l'autre peuvent partager une représentation fausse, et leur relation en souffrira sans qu'aucun des deux ait quoi que ce soit à se reprocher. La décision qui change est de savoir où chercher quand quelque chose ne va pas : dans la sincérité de l'autre, ou dans les faits. C'est une distinction qui concerne aussi l'instrument, car la méta-interaction sert à la première et non à la seconde : elle reconstruit ce que chacun croyait, non ce qui était vrai.

Manipulation — Figure dans laquelle un agent obtient d'un autre un comportement en agissant sur son processus décisionnel plutôt que sur ses raisons, et de telle manière que l'autre croie avoir choisi. Elle n'exige pas le mensonge : on manipule aussi avec des affirmations vraies, en choisissant quoi dire, quand le dire et dans quel ordre, et en exploitant ce que l'on sait des programmes de l'autre. Elle se distingue du commandement, qui est déclaré — celui qui le reçoit sait qu'on lui a demandé quelque chose — et du pouvoir, qui l'emporte sur une résistance au grand jour : la manipulation est un pouvoir qui se cache dans le choix de l'autre. Dans les termes du chapitre précédent elle opère en produisant

chez celui qui la subit la reconnaissance d'une figure qui convient à celui qui l'exerce. Le désalignement est maximal et différé : l'un reconnaît une persuasion ou une aide, l'autre reconnaît une manipulation seulement après, et souvent jamais. La décision qui change est de savoir si l'on doit répondre à l'autre ou s'y soustraire : à un argument on répond, à une manipulation non.

Confrontation et évitement — Figure dans laquelle les deux peuvent, ou ne peuvent pas, parler de ce qui se passe entre eux. Elle ne dépend pas des autres figures : on évite la confrontation dans des relations coopératives et on l'affronte dans des relations hostiles. C'est la figure qui ouvre ou ferme l'accès à l'instrument que cet essai propose, car une méta-interaction n'est possible que là où la confrontation l'est ; et ce qui la rend impossible est presque toujours une prévision — qu'en parler aggraverait les choses — dont la vérité n'est jamais mise à l'épreuve, car éviter la confrontation empêche aussi de découvrir qu'il en serait allé autrement. Le désalignement typique est que l'un des deux tienne la question pour réglée et l'autre tienne qu'on n'en parle jamais. La décision qui change est de savoir s'il vaut la peine de soulever une question ou s'il convient d'agir par d'autres voies.

Les figures relationnelles d'asymétrie

Dépendance, indépendance — Un agent dépend d'un autre dans la mesure où la perte de la relation lui coûte : parce qu'il en reçoit des ressources, des informations ou des services qu'il ne trouve pas ailleurs, ou parce qu'il ne dispose pas d'alternatives. La dépendance est rarement égale, et elle est le fondement du pouvoir plus qu'une conséquence de celui-ci : l'un peut l'emporter sur les préférences de l'autre dans la mesure où interrompre la relation coûte davantage à l'autre. Elle s'applique à tout agent qui reçoit quelque chose d'un autre, et décrit un fournisseur et un client comme une cellule et le tissu qui la nourrit. La décision qui change est de savoir jusqu'où l'on peut insister, et s'il convient de se procurer une alternative : qui réduit sa propre dépendance modifie la figure de pouvoir sans avoir agi directement sur elle.

Pouvoir exercé, pouvoir subi — Le pouvoir est défini au chapitre *Définitions* comme la capacité de modifier de manière significative et durable l'état ou le comportement d'un autre agent, même contre ses préférences. Comme figure relationnelle il doit être reconnu des deux côtés, et les deux côtés s'accordent rarement : celui qui l'exerce ne s'en aperçoit presque jamais dans la juste mesure, parce qu'il voit ses propres actions comme normales et n'a pas accès à l'effort qu'il en coûte de les subir ; celui qui le subit tend à lui attribuer une intention plus précise qu'elle ne l'est. La décision qui change est de savoir où chercher la cause d'un comportement propre que l'on ne s'explique pas : dans sa propre volonté, ou dans la marge que quelqu'un d'autre a rétrécie.

Commandement, obéissance — Le commandement est défini au chapitre *Définitions* comme la communication dont l'objectif est que le destinataire exécute une action. Il y a figure de commandement et d'obéissance lorsque l'exécution est attendue de manière stable, et c'est ici qu'il faut introduire l'autorité, que la notion de pouvoir ne couvre pas : l'autorité est le droit de commander, concédé par celui qui obéit ou établi par un tiers, tandis que le pouvoir l'emporte sur une résistance sans avoir besoin d'aucun droit. La différence se voit lorsque l'autorité disparaît et que le pouvoir demeure, ou l'inverse. La décision qui change est de savoir si, devant un commandement, la question est d'obéir ou de négocier : contre le pouvoir on ne peut que résister ou se soustraire, tandis que l'autorité peut être contestée en invoquant la source même qui l'a instituée.

Supériorité et infériorité reconnues — C'est le jugement comparatif que chacun des deux porte sur la valeur de l'autre, et il n'a de sens que si l'on précise par rapport à quoi : compétence dans un domaine déterminé, mérite, capacité d'obtenir des résultats. Reconnues signifie qu'il s'agit d'un jugement et non d'un fait : les deux peuvent en porter d'opposés, et en porter de différents domaines par domaine. Entre organisations le jugement prend le nom de réputation et a le même effet : il détermine de qui l'on accepte normes et procédures. La décision qui change est de savoir à qui demander et de qui accepter une correction.

Supériorité et infériorité institutionnelles — C'est la position établie par un tiers — une organisation, une communauté, une norme — indépendamment du jugement des deux. Elle se distingue de la précédente parce qu'elle n'a pas besoin d'être reconnue pour opérer, et la divergence entre les deux est une source typique de conflit : le cas ordinaire est celui de qui est supérieur en grade et inférieur en compétence aux yeux de son subordonné. La décision qui change est de savoir par quelle voie l'on peut obtenir quelque chose : vers une supériorité institutionnelle on passe par la procédure, vers une supériorité reconnue on passe par l'argument.

Donner et recevoir un service — Un agent accomplit pour un autre une action que celui-ci n'accomplirait pas de lui-même, ou pas aussi bien. C'est la figure déjà décrite au chapitre *Concepts fondamentaux* entre un sous-agent et son agent supérieur, mais elle se rencontre aussi entre pairs et entre agents de toute espèce : un organe pour l'organisme, un fournisseur pour un client, un programme pour son utilisateur. Le désalignement typique est que le service donné soit reçu comme un commandement ou comme un contrôle, et la raison en est que tout service comporte de décider, à la place d'un autre, une part de ce qui le concerne. La décision qui change est de savoir si une offre d'aide doit être acceptée ou refusée, et la question utile n'est pas de savoir si elle est bien intentionnée — elle l'est d'ordinaire — mais combien de liberté elle retire à celui qui la reçoit.

Dettes et gratitude — Figure dans laquelle ce qu'un agent a reçu d'un autre laisse une obligation. Elle est à *Donner et recevoir un service* ce que la relation est à l'interaction : l'acte et son sédiment. Elle est asymétrique d'une manière caractéristique : celui qui a donné se souvient plus longtemps et estime plus haut que celui qui a reçu, et parfois celui qui reçoit se trouve avec une dette que le donneur n'entendait pas créer — c'est la manière ordinaire dont un don devient un poids. La décision qui change est de savoir si l'on doit quelque chose, et donc si une demande ultérieure est une demande ou un recouvrement.

Les figures relationnelles quant à la valence et aux effets

Attraction, répulsion — Le chapitre *Les émotions* les définit comme les deux valences avec lesquelles une émotion parvient à la conscience. Lorsque l'objet de l'émotion est un autre agent, elles constituent une figure relationnelle, et c'est la seule de la liste qu'il n'y ait pas à reconstruire : elle se sent. Elle exige toutefois un agent doté de conscience, et aux autres elle ne s'applique pas. Comme toute émotion elle peut être mixte, et la coprésence d'attraction et de répulsion envers la même personne est un conflit intra-agent, non une incertitude à résoudre. La décision qui change est de savoir quel poids donner à sa propre orientation affective dans les choix concernant cet agent — ce qui est une décision et non une donnée : on choisit de travailler avec ceux qui ne plaisent pas, et l'on évite ceux qui plaisent.

Confiance, méfiance — C'est la fiabilité que l'on attribue au comportement futur de l'autre : dans quelle mesure on s'attend à ce qu'il fasse ce qu'il a dit, et à ce qu'il n'abuse pas de ce qu'on lui a concédé. Elle doit être tenue distincte de l'attraction, avec laquelle il est facile de la confondre : on fait confiance à des personnes qui ne plaisent pas, et l'on ne fait pas confiance à des personnes qui plaisent beaucoup. Elle présuppose un certain degré de connaissance de l'autre, et elle est asymétrique par construction, parce que chacun a de soi une évidence qu'il n'a pas de l'autre. La décision qui change est de savoir jusqu'où s'exposer : quoi confier, combien vérifier, quelles garanties demander. Entre organisations, c'est l'entrée qui décide si un contrat doit tout prévoir ou peut laisser quelque chose d'ouvert.

Domage, offense — Un dommage est un effet défavorable produit chez un agent par les actions d'un autre. Une offense est un dommage auquel celui qui le subit attribue une intention. La distinction est celle que le chapitre *Définitions* fixe à propos de la responsabilité, et c'est ici que le désalignement atteint son maximum : celui qui a produit le dommage souvent ne s'en est pas aperçu, et le juge presque toujours non intentionnel, parce qu'il a accès à ses propres intentions ; celui qui le subit n'a accès qu'à l'effet, et pour lui l'attribution d'une intention est une hypothèse que rien dans l'expérience immédiate ne dément. La décision qui change est de savoir quelle réparation demander et laquelle

offrir : à un dommage on répond en le réparant, à une offense non, et c'est la raison pour laquelle tant de réparations matérielles ne closent rien. Ce qu'un dommage ou une offense déposent dans le temps est le ressentiment, et son interruption délibérée est le pardon : ce ne sont pas des figures à part mais le sédiment de celle-ci, et la décision qu'ils changent est de savoir si le compte est clos — tant qu'il reste ouvert, chaque interaction suivante s'y ajoute.

Justice, injustice — Figure dans laquelle chacun des deux mesure ce qui se passe entre eux contre un critère qu'il suppose commun : la proportion entre ce que chacun donne et reçoit, la correspondance entre ce que l'on a fait et ce que l'on obtient, ou l'égale application d'une règle aux deux — cette dernière est la figure que l'on reconnaît quand l'autre se pardonne ce qu'il nous reproche. Elle présuppose le partage des critères de jugement, au sens de l'entrée *Partage* : sans un critère supposé commun il n'y a pas d'injustice mais seulement un désavantage. Elle doit être tenue distincte de *Dommage, offense* : on peut subir une injustice sans qu'aucun dommage n'ait été voulu, et même sans qu'aucun dommage n'ait été fait, comme dans la répartition inégale d'un avantage. Le désalignement y est la règle et non l'exception, car chacun applique le critère depuis sa propre position et se souvient de ses propres contributions mieux que de celles de l'autre : deux agents peuvent se sentir tous deux lésés, sincèrement et au même moment. La décision qui change est la plus grosse de la liste, car le sentiment d'injustice transforme un désavantage en une prétention : qui se juge en perte se demande s'il convient de changer, qui se juge lésé se demande s'il faut exiger, et s'autorise des actions qu'il s'interdirait autrement. L'essai ne dit pas ce qui est juste : il décrit la figure, c'est-à-dire le fait que chacun mesure contre un critère et que les critères peuvent diverger.

La sélection

Choisir et écarter, être choisi et être écarté — Un agent cherche, maintient, relâche ou interrompt une relation, et change son choix au cours du temps. Les positions sont quatre et doivent être tenues distinctes, car chacune change des choses différentes : choisir, écarter, être choisi, être écarté. C'est une décision plus qu'une figure relationnelle, mais elle est reconnue comme figure des quatre côtés, et elle est asymétrique comme aucune autre entrée, car le choix peut ne pas être réciproque. Elle est toujours relative à un rôle déterminé : on choisit quelqu'un comme collègue et non comme ami, et confondre les deux plans est une manière ordinaire de blesser sans le vouloir. De toutes les figures c'est celle qui a les plus grandes conséquences : pour un être humain, être écarté est l'une des reconnaissances les plus chargées, et le signal émotionnel qui l'accompagne a l'intensité que le chapitre *Les émotions* attribue aux messages que la conscience ne peut pas ignorer.

La sélection est le point où toutes les entrées précédentes confluent, et c'est la raison pour laquelle cet essai les traite comme des entrées d'un processus décisionnel. Aucune d'elles ne la détermine. L'attraction l'oriente dans la plupart des cas, mais pas nécessairement ; la dépendance peut maintenir en vie une relation que toutes les autres entrées déconseilleraient ; une supériorité institutionnelle peut soustraire entièrement le choix, car un subordonné ne choisit pas son supérieur. Ce que la méthode agentologique peut faire, c'est rendre explicites les termes sur lesquels la décision se prend, et en particulier distinguer ce qui appartient à la situation relationnelle de ce qui appartient seulement à la manière dont on l'a perçue.

Le troisième agent

Jusqu'ici les figures relationnelles ont été considérées à l'intérieur d'un couple. Mais un couple est rarement seul : autour de deux agents il y en a d'autres, et chacun d'eux, en entrant dans le tableau, modifie la figure entre les deux. Toute figure duelle est immergée dans autant de triades qu'il y a de tiers qui comptent, et le principe est que la figure entre A et B dépend aussi des figures que A et B ont avec C.

La triade comme unité de l'analyse a un précédent connu : la théorie de l'équilibre cognitif de Fritz Heider (*Attitudes and Cognitive Organization*, 1946), qui considère les trois relations d'une triade comme représentées par l'un des trois, et prévoit une pression vers les configurations où leurs signes

s'accordent. De Heider je prends l'idée que la triade est perçue par quelqu'un et décrite par personne ; je ne prends pas le calcul des signes, parce qu'il présuppose des figures dotées d'un signe et symétriques, et que la plupart des entrées énumérées plus haut — commander, dépendre, donner un service — ne sont ni l'un ni l'autre.

Le cas le plus simple se dérive de ce que l'essai possède déjà, sans rien ajouter. Si A veut commander B, et que C veut aussi commander B, A et C sont en compétition : l'obtention de l'objectif de l'un réduit la probabilité que l'autre obtienne le sien, ce qui est la définition donnée au chapitre *Définitions*. Aucun des deux n'a rien fait à l'autre, et ils peuvent ne s'être jamais rencontrés : la figure entre eux naît des figures que chacun a avec un tiers.

La règle ne vaut cependant pas pour toute figure, et la condition est celle déjà indiquée à l'entrée Partage : elle dépend de ce que l'on veut du tiers admettre ou non deux destinataires. Si A et C voulaient tous deux coopérer avec B, ils ne seraient pas pour autant en compétition, car la coopération de B est divisible ; le commander ne l'est pas, ou ne l'est que mal, et de fait qui sert deux maîtres les sert mal tous les deux. La question à se poser devant une triade n'est donc pas de savoir si les deux ont la même figure avec le tiers, mais si ce qu'ils veulent de lui est divisible.

Autres cas récurrents, à titre d'exemple. Si A dépend de B et que B dépend de C, A dépend de C sans le savoir, et ses prévisions sur le comportement de B resteront fausses tant qu'il ne s'en apercevra pas. Si A et B obéissent tous deux à C, ils se trouvent dans la même position à son égard, et cela abaisse le coût de coopérer entre eux : c'est la manière ordinaire dont naissent les alliances entre subordonnés. Si A imite C et que B imite C également, A et B se trouveront semblables sans s'être cherchés, et la ressemblance produira compétition ou reconnaissance selon que les ressources dont elle les rapproche sont divisibles ou non. Les combinaisons possibles se comptent par centaines — vingt-deux entrées pour chacun des trois côtés — et ne s'énumèrent pas : ce que l'on peut donner, c'est le principe et la manière de l'appliquer cas par cas, ce qui est ce qu'offre la méthode agentologique en général.

Reste enfin, et ici plus fortement qu'ailleurs, ce que le chapitre précédent a établi : la triade est perçue. Celle de A est celle de A, et peut ne correspondre à aucune autre. C peut ignorer que A existe, et se trouver entravé par quelqu'un dont il ne sait rien ; A peut attribuer à B une figure avec C que B n'a pas ; deux agents peuvent agir l'un contre l'autre avec détermination, chacun à l'intérieur d'une triade que l'autre ne reconnaîtrait pas. La reconnaissance du tiers n'est pas plus fiable que celle du second, et elle est même moins vérifiable, parce que le tiers n'est d'ordinaire pas présent pour être interrogé.

Le désalignement

Le rendement de ce chapitre ne tient pas à la liste, qui est ordinaire, mais à ce que l'on obtient en la parcourant deux fois : une fois de son propre point de vue, une fois de celui de l'autre. Les cas où les deux reconnaissances divergent ont une forme récurrente. L'un des deux rend un service, l'autre reçoit un commandement : c'est la mère qui téléphone trois fois par jour par laquelle la *Préface* a commencé, et la raison pour laquelle la dire envahissante n'explique rien. L'un coopère, l'autre se sent exploité, parce que l'objectif commun dont le premier parle est l'objectif du premier. L'un imite par reconnaissance, l'autre subit une appropriation. L'un contrôle par souci, l'autre reçoit de la méfiance : c'est le cas de Giulia et de Marco, où la divergence ne portait pas sur les faits, que tous deux connaissaient, mais sur les figures relationnelles que chacun y reconnaissait.

Dans tous ces cas la reconnaissance des deux est incompatible, et aucun des deux ne ment. Il n'y a rien, dans l'expérience de chacun, qui puisse la corriger : chacun perçoit ses propres intentions et les effets qu'il subit, non les intentions d'autrui ni les effets qu'il produit. Il faut une opération qui confronte les deux reconnaissances, et c'est celle que le chapitre *Application aux relations sociales* appelle méta-interaction. Le présent chapitre en précise l'objet : ce qu'une méta-interaction confronte, ce ne sont pas les faits, sur lesquels d'ordinaire il n'y a pas de désaccord, mais les figures, que les faits ne suffisent pas à établir. Avec deux exceptions, que la liste contient désormais : dans le mensonge et dans la manipulation la divergence n'est pas l'erreur des deux mais le résultat voulu par l'un, et le

remède n'est pas de reconstruire ensemble ce que chacun croyait, mais de cesser de tenir pour bon ce qui est dit.

Application aux relations sociales

Le chapitre précédent a appliqué l'appareil conceptuel agentologique à la vie interne d'une personne. Ce chapitre l'applique à ce qui advient entre les personnes, et introduit un concept que le noyau axiomatique ne contient pas — le bien-être lié à la qualité des relations.

Ce que dit la Psychologie de l'interagir

Cet essai reprend des idées formulées dans un essai précédent, *Psychologie de l'interagir*, consacré spécifiquement aux interactions humaines. Cet essai soutient qu'une bonne part de nos joies et de nos peines dépend de la qualité des relations sociales, lesquelles dépendent à leur tour de la qualité des interactions quotidiennes qui les construisent. C'est une thèse largement soutenue par la recherche sur le bien-être subjectif, mais elle doit être entendue comme l'indication d'un facteur pertinent parmi d'autres — personnalité, conditions matérielles, santé — plutôt que comme une hiérarchie rigide entre les causes du bonheur.

Psychologie de l'interagir a pour objet l'interagir : non la personnalité, non la communication, non les relations prises isolément, mais le processus par lequel deux personnes ou plus, à travers une succession de transactions, modifient, maintiennent ou rétablissent une situation relationnelle — la configuration de rôles, d'attentes réciproques, de rapports de force, de contraintes et de climat émotionnel qui se présente entre elles à un moment donné. Sa thèse centrale est que les êtres humains ne réagissent pas à la situation relationnelle, mais à l'interprétation qu'ils en construisent. Le chapitre *Définitions* reprend le terme en un sens plus étroit — la configuration des objectifs, des programmes réciproquement pertinents, de l'asymétrie et des effets réciproques — laissant de côté les rôles, les attentes et le climat émotionnel que *Psychologie de l'interagir* y comprend.

L'interprétation de la situation rencontre le système motivationnel de la personne — besoins, désirs et valeurs — et il en naît un objectif. L'objectif appelle une logique psychologique, c'est-à-dire un ensemble relativement stable de règles du type «si je me trouve dans une situation de ce genre et que je veux obtenir ce résultat, alors il convient de se comporter de telle manière». La logique oriente une stratégie, la stratégie se traduit en transactions, et chaque transaction modifie la situation relationnelle, à partir de laquelle le tour recommence.

Cet essai distingue quatre niveaux : la transaction, unité élémentaire de l'interagir et comportement perceptible par l'autre ; l'interaction, c'est-à-dire la succession de transactions à l'intérieur d'une même situation relationnelle ; la relation, c'est-à-dire le lien qui se sédimente dans le temps entre des individus déterminés ; le réseau social, c'est-à-dire l'ensemble des relations dans lesquelles une personne est insérée. Et il observe qu'une situation relationnelle peut être aussi intériorisée : l'autre — un parent disparu, une personne lointaine, une figure imaginée — ne produit pas de transactions réelles, mais il est représenté intérieurement et oriente tout autant le comportement de celui qui l'intériorise.

Ce qu'ajoute l'agentologie

L'appareil conceptuel de *Psychologie de l'interagir* — situation relationnelle, interprétation, motivation, logique psychologique, stratégie, transaction — peut être relu comme une application du vocabulaire agentologique aux interactions entre personnes. Cette relecture n'est pas un exercice de traduction : sur quatre points elle met en lumière quelque chose qui, dans l'essai précédent, restait implicite.

Que l'on réagisse à l'interprétation et non à la situation correspond à l'axiome A5 uni à la définition de la perception : le comportement d'un agent dépend de la situation perçue, et non de la situation, pour la raison qu'à la situation aucun agent n'a accès — il a accès en revanche aux informations que sa perception lui livre. Le gain n'est pas de généralité abstraite : si la divergence entre la situation et la

situation perçue est une condition de tout agent, alors c'est un lieu où intervenir même là où l'on n'a pas l'habitude de le chercher, dans une organisation qui ne voit pas son propre marché comme dans un programme qui ne voit pas le cas pour lequel il n'a pas été écrit.

La logique psychologique est, dans le vocabulaire agentologique, un programme : une structure interne qui contraint les actions qu'un agent peut produire à partir d'une perception donnée. *Psychologie de l'interagir* observe que les logiques ressemblent au logiciel par leur structure conditionnelle, mais s'en distinguent parce qu'elles se modifient de l'intérieur, sous l'effet de l'expérience même qu'elles contribuent à orienter. En agentologie cette différence ne passe plus entre l'homme et la machine : c'est la distinction, fixée par l'axiome A6, entre programmes plastiques et programmes rigides, qui vaut pour n'importe quel agent. Il y a des programmes rigides chez les êtres humains — et c'est sur cela que le chapitre suivant construit la notion de changement thérapeutique — et des programmes plastiques dans les machines.

La place que cet essai laisse déclarément vide, celle des émotions, est occupée par le chapitre *Les émotions* du présent essai. Le vécu immédiat d'attraction et de répulsion n'y est pas une donnée à enregistrer : c'est un message qu'un agent inconscient envoie à la conscience, avec un objet, une valence et une intensité. Et la coprésence d'attraction et de répulsion envers la même personne, que *Psychologie de l'interagir* constate sans l'expliquer, devient la réception simultanée de deux messages provenant d'expéditeurs différents, c'est-à-dire un conflit intra-agent. Cet essai dit que le vécu est réellement double ; l'agentologie dit pourquoi il l'est.

Il arrive que l'on se comporte en fonction de quelqu'un qui n'est pas présent : un parent mort, une personne lointaine, un juge imaginaire. *Psychologie de l'interagir* enregistre le phénomène et le traite comme un cas particulier de situation relationnelle, en observant qu'à la différence des autres il n'est pas vérifiable, parce que l'autre n'est pas là pour nous démentir. L'agentologie le décrit autrement : ce quelqu'un n'est pas une situation, c'est un agent — un sous-agent de la personne, construit avec les interactions qu'elle a réellement eues avec l'original, et qui continue d'évaluer, d'approuver et de reprocher selon les critères qui lui ont été attribués. La différence n'est pas de mots. Un sous-agent peut être modifié, et on le modifie sans que la personne réelle en sache rien ; et il continue d'opérer après la mort de cette personne. C'est la raison pour laquelle une thérapie peut s'occuper du rapport avec un défunt, et le chapitre suivant y revient.

Enfin le réseau social, qui est dans cet essai l'ensemble des relations dans lesquelles un individu est inséré. L'agentologie y ajoute une question : celle de savoir si le réseau est lui-même un agent d'ordre supérieur. Selon l'axiome A8 il l'est lorsque, au niveau du réseau, émergent des perceptions et des décisions non réductibles à la somme de celles de ses membres — une procédure de décision commune, une frontière reconnaissable par rapport à l'extérieur, une persistance distincte de celle des membres pris un à un. Une famille peut satisfaire ces conditions ; un groupe de connaissances presque jamais. C'est la distinction qui permet d'établir quand il a un sens d'attribuer un objectif à un groupe, et quand ce n'est au contraire qu'une façon de parler.

Ces correspondances n'enlèvent rien à la spécificité de *Psychologie de l'interagir* : la phénoménologie du vécu immédiat, la distinction entre besoins, désirs et valeurs, les configurations triadiques plus ou moins équilibrées restent des contenus propres à cet essai, que l'agentologie ne prétend pas corriger. Ce que l'agentologie ajoute, c'est un langage capable de situer ce modèle, avec d'autres, à l'intérieur d'un appareil conceptuel plus général, applicable aussi à des organisations et à des agents artificiels.

La méta-interaction

L'essai *Psychologie de l'interagir* propose la méta-interaction : une interaction qui a pour objet une autre interaction. L'objectif est de reconstruire ensemble les interprétations, les motivations et les logiques qui ont engendré une interaction précédente, ou de négocier une interaction future.

Dans le présent essai la méta-interaction n'est plus seulement une pratique conçue pour les relations humaines : c'est un cas particulier d'une opération générale — la communication réflexive sur une interaction passée ou possible dans le futur — applicable en principe à tout couple ou groupe d'agents, humains ou artificiels, capables de communiquer entre eux et d'un accès, si partiel soit-il, à leur propre état interne passé (mémoire). Si une bonne part du bien-être individuel dépend de la qualité des relations, et si la qualité d'une relation dépend de la qualité des interactions qui la construisent, la méta-interaction — sous la forme générale ici proposée — est un instrument pour intervenir délibérément sur cette qualité.

Quels sont les désaccords qui la rendent nécessaire, la section *Le désalignement* du chapitre *Répertoire des figures relationnelles* le dit : ils ne portent pas sur les faits, sur lesquels d'ordinaire il n'y a pas de désaccord, mais sur les figures que chacun y reconnaît.

Application à la psychothérapie et à l'autothérapie

Le chapitre précédent a appliqué l'agentologie aux rapports entre personnes. Ce chapitre applique le même appareil conceptuel aux rapports entre les programmes internes à une même personne, afin de faciliter la compréhension — et, quand c'est possible, le dépassement — des conflits motivationnels.

Conflit intra-agent

Conflit intra-agent — Cas de conflit où les parties en contraste ne sont pas des agents distincts, mais des composants d'un même agent : deux programmes ou plus, ou deux sous-agents ou plus, qui, activés par la même perception évaluée de manières différentes, tendent vers des actions antagonistes ou incompatibles entre elles.

Les niveaux du conflit

Le conflit se présente à trois niveaux, qui se distinguent par l'endroit où tombe la frontière entre les parties en contraste.

Le premier niveau est le conflit entre agents distincts : deux personnes, mais aussi deux organisations, ou une personne et une institution. C'est le niveau dont se sont occupés les chapitres *Répertoire des figures relationnelles* et *Application aux relations sociales*, sous les noms de compétition, de pouvoir, de dépendance et d'interaction mixte.

Le deuxième est le conflit entre sous-agents d'une même personne. Parmi ceux-ci il faut compter les modèles intériorisés d'autres personnes. Un parent intériorisé est un sous-agent construit avec les interactions que l'on a eues avec lui, qui continue d'évaluer, d'approuver et de reprocher selon des critères qui lui ont été attribués. Il en découle deux conséquences qui comptent pour une thérapie : ce sous-agent peut être modifié sans que la personne réelle change, et il continue d'opérer après la mort ou l'éloignement de l'original.

Le troisième est le conflit entre programmes d'un même agent qui n'ont pas l'autonomie suffisante pour qu'on les appelle sous-agents : une habitude contre un propos, un réflexe contre une règle apprise. La frontière entre ce niveau et le précédent est celle qui a été indiquée au chapitre *Concepts fondamentaux*, et elle est graduelle : dans bien des cas elle n'est pas décidable, et aux fins d'une thérapie il peut n'être pas nécessaire de la décider, car dans les deux cas il s'agit de modifier des programmes. La distinction compte toutefois pour la difficulté de l'entreprise : modifier deux programmes qui résident dans le même agent est vraisemblablement plus facile que d'en modifier deux qui résident dans des sous-agents différents, précisément parce que ces derniers sont plus autonomes.

Les trois niveaux peuvent s'influencer mutuellement. Un conflit entre personnes peut être soutenu, chez chacune d'elles, par un conflit interne qui la porte à interpréter l'autre d'une certaine manière et non d'une autre. Un sous-agent intériorisé est le sédiment de conflits interpersonnels passés. Et un conflit interne peut se décharger vers l'extérieur, en cherchant au-dehors un adversaire plus facile à combattre que celui qu'on a au-dedans.

L'instrument thérapeutique pour aborder le premier niveau a déjà été défini : la méta-interaction, c'est-à-dire la reconstruction conjointe de ce que les participants ont perçu, voulu et décidé dans une interaction passée. Pour le deuxième et le troisième niveau valent les procédés décrits dans les sections qui suivent.

Une précaution, enfin. Le niveau auquel un conflit appartient n'est pas toujours évident : il est lui-même l'objet d'une reconstruction, et donc d'une erreur possible. Attribuer à un contraste interne ce qui naît d'un contraste réel avec une autre personne, ou l'inverse, n'est pas une éventualité rare : c'est l'une des formes les plus communes d'erreur dans l'analyse de soi.

Pouvoir intra-agent et acrasie

Pouvoir intra-agent — Un programme exerce un pouvoir sur un autre programme du même agent lorsqu'il l'emporte de manière significative et durable sur les actions que le second aurait autrement produites, même lorsque cela contrarie l'objectif vers lequel le second tend. C'est le cas particulier du pouvoir où celui qui l'exerce et celui qui le subit ne sont pas des agents distincts, mais des composants d'un même agent : deux programmes, ou deux sous-agents.

Une personne sait ce qu'elle veut faire, et fait le contraire. Aristote appelait cela acrasie ; la philosophie contemporaine l'appelle faiblesse de la volonté. Dans les termes de cet essai les deux pôles sont deux programmes : l'un poursuit l'objectif que la personne, en y réfléchissant, déclare être le sien ; l'autre est une impulsion, une habitude, une compulsion qui va en sens contraire. C'est le second qui l'emporte, et qui l'emporte de manière stable. Personne ne la contraint de l'extérieur : la contrainte vient du dedans.

Que cela soit possible ne contredit pas l'axiome selon lequel tout agent peut se modifier lui-même. Cet axiome garantit la capacité d'agir sur soi, non la cohérence de ce qui en résulte : un agent peut travailler pour soi et contre soi.

Le chapitre *Les émotions* a indiqué le canal par lequel ce pouvoir s'exerce : le signal émotionnel, et la promesse de plaisir ou de douleur qui l'accompagne.

Le changement thérapeutique comme apprentissage dirigé

Le changement que vise une psychothérapie peut être relu comme un apprentissage intentionnel : la modification d'un ou de plusieurs programmes à la suite de l'expérience que constitue le travail thérapeutique lui-même.

Il va de soi que les programmes à corriger dont nous parlons ne peuvent pas être lus comme ceux d'un ordinateur. On peut toutefois chercher à les reconstruire à partir des comportements et des émotions des sujets : c'est une activité d'abduction qu'en informatique on appelle *reverse engineering*.

Un programme peut être plastique ou rigide ; la thérapie travaille sur les programmes plastiques, et une part non négligeable du travail thérapeutique consiste précisément à établir quels programmes sont, à un moment donné, effectivement plastiques et lesquels ne le sont pas.

Ce qui distingue cet apprentissage d'un apprentissage quelconque, c'est son point de départ : non pas l'acquisition d'une information nouvelle sur l'environnement extérieur, mais une perception plus exacte d'un conflit que le programme précédent impliquait. Rendre explicite un conflit intra-agent — reconnaître quels sous-agents y sont impliqués et quels objectifs ils poursuivent — est souvent, en soi, la condition qui rend de nouveau plastique un programme autrement rigide. Ce n'est pas une conséquence de l'appareil conceptuel : c'est une hypothèse que l'appareil permet de formuler avec précision, et que seul le travail clinique peut confirmer ou démentir.

La méthode appliquée à soi-même

Les questions du questionnaire agentologique peuvent être formulées à la première personne, et adressées non pas à un phénomène extérieur mais à ses propres agents et sous-agents : lesquels d'entre eux sont impliqués dans ce conflit ? Quelle est la situation qui l'active ? Que perçoit chacun d'eux ? Quelles figures chacun d'eux reconnaît-il entre lui et les autres ? Quels programmes, chez chacun, se révèlent pertinents ? Quel objectif poursuit chacun ? Quelle action tend-il à produire ? Comment ces actions, en compétition entre elles, modifient-elles mon comportement effectif ? Quelle nouvelle situation interne ou externe en résulte-t-il ?

Appliqué à soi-même, cet exercice ne requiert pas nécessairement un contexte clinique : il peut être conduit par écrit, comme pratique réflexive ordinaire.

Le problème spécifique de l'autothérapie

L'application à soi-même rencontre cependant une difficulté que l'application à un phénomène extérieur ne rencontre pas : l'observateur et l'observé coïncident. La perception sur laquelle l'auto-analyse se fonde est produite par les programmes mêmes que l'auto-analyse voudrait examiner, et elle est donc susceptible d'être déformée aux points les plus pertinents. Cet essai a déjà rencontré le problème, à propos de la liberté : la divergence entre les actions qu'un agent perçoit comme disponibles et celles qui le sont objectivement est le cas typique de l'auto-illusion. Dans l'autothérapie ce risque n'est pas une exception, mais la condition ordinaire de départ.

Certaines précautions peuvent atténuer le problème, sans l'éliminer : l'externalisation par l'écriture, qui fait de sa propre reconstruction un objet stable que l'on peut relire et remettre en question à distance de temps ; la relecture différée, qui permet de confronter la reconstruction faite à chaud avec une lecture plus détachée ; la confrontation avec un tiers — un professionnel ou, avec les limites que cela comporte, une personne de confiance — qui ne partage pas les mêmes programmes et peut donc percevoir ce que l'agent, seul, ne perçoit pas systématiquement.

Avertissement sur l'usage de l'agentologie comme instrument thérapeutique

L'agentologie offre un langage pour décrire des conflits motivationnels ordinaires et pour y réfléchir — l'ambivalence, la procrastination, les schémas relationnels dysfonctionnels d'intensité modérée. Elle n'offre pas un instrument diagnostique ou thérapeutique pour des conditions qui requièrent un traitement clinique — traumatisme, psychose, dépression majeure, troubles qui compromettent la capacité même de percevoir avec exactitude sa propre situation. En ces cas la médiation d'un professionnel n'est pas une option parmi d'autres, mais la condition pour qu'une analyse quelconque, agentologique ou non, puisse être dite fiable.

Un exemple d'application

Les deux chapitres précédents ont introduit des concepts applicatifs de manière abstraite. Ce chapitre les met à l'épreuve sur un exemple imaginaire, qui ne démontre rien mais permet de vérifier sur un cas ce qui a été affirmé dans la *Préface*. Le critère est celui-ci : une distinction vaut ce qu'elle coûte lorsque quelque chose change — soit l'action que l'on choisit, soit ce que l'on s'attend à voir advenir. Dans l'histoire qui suit les distinctions introduites changent l'une et l'autre chose, et à la fin du chapitre il est dit où.

Marco travaille depuis trois ans dans le service dirigé par Giulia. Depuis quelques mois il remet systématiquement en retard les rapports hebdomadaires, bien qu'il sache que le retard irrite Giulia et pèse sur son évaluation annuelle. Giulia, de son côté, a réagi aux retards en intensifiant les contrôles et en demandant des mises à jour quotidiennes — ce que Marco vit comme de la méfiance, et qui renforce, au lieu de l'atténuer, sa tendance à remettre à plus tard. La relation entre les deux, construite en trois ans de collaboration surtout coopérative, s'est progressivement chargée d'éléments spécifiques de signe négatif.

En appliquant la méthode agentologique : les agents impliqués sont Marco et Giulia — le service dans son ensemble ne satisfait pas, en ce cas, le critère de A8, car il n'existe pas de procédure décisionnelle impliquant le service comme tel, indépendante des personnes prises une à une, et il reste donc une pluralité, non un agent d'ordre supérieur. La situation initiale est une relation sédimentée sur une histoire de collaboration, mais avec une interprétation récente asymétrique : Marco perçoit les contrôles de Giulia comme un jugement sur sa propre fiabilité ; Giulia perçoit les retards de Marco comme un signe de désengagement. Dans les termes du chapitre *Les figures relationnelles*, les deux reconnaissent des figures différentes dans la même interaction : Giulia se voit du côté de qui rend un service — attention, soutien, vérification — et Marco reçoit un commandement ; et c'est la divergence entre les deux lectures, et non l'un ou l'autre pris isolément, qui explique la détérioration.

Les programmes pertinents chez Marco comprennent aussi bien un objectif professionnel explicite — conserver une bonne évaluation — qu'une habitude d'évitement qui s'active face à des tâches perçues comme sources de jugement négatif : les deux programmes sont en conflit, et l'habitude d'évitement exerce de fait un pouvoir sur le programme professionnel, l'emportant sur lui bien que l'objectif déclaré de Marco reste de remettre son travail à temps — un cas d'acrasie au sens ici défini. L'action observable — le retard répété — est le résultat d'un conflit intra-agent qui reste implicite tant que personne ne le nomme.

Du côté de Giulia, l'action d'intensifier les contrôles est un exercice de pouvoir qui, rencontrant l'interprétation de Marco, produit l'effet opposé à celui qui était visé : au lieu de rétablir la fiabilité, il renforce le programme d'évitement, parce qu'il confirme à Marco l'interprétation qu'il redoute. L'interaction entre les deux, bien qu'enracinée dans une relation à base coopérative, prend localement une dynamique compétitive — exactement le cas d'interaction mixte déjà prévu dans la remarque sur la coopération et la compétition.

Le tournant arrive lorsque Giulia propose une confrontation directe sur ce qui est en train de se passer entre eux : une méta-interaction, au sens défini au chapitre *Application aux relations sociales*. Au cours de cette confrontation il apparaît que Marco n'avait jamais perçu, parmi les actions dont il disposait, celle d'expliquer ouvertement les difficultés de charge de travail qui étaient à l'origine des retards : non parce qu'elle lui était objectivement interdite, mais parce qu'elle n'entraînait pas dans sa liberté, au sens ici défini — un cas, précisément, de divergence entre les actions qu'un agent tient pour disponibles et celles qui le sont objectivement pour lui.

Le résultat immédiat de la confrontation est un changement de perception, non encore un changement de programme : Marco réinterprète les contrôles de Giulia comme du souci et non comme de la méfiance, et Giulia réinterprète les retards comme un signal de surcharge et non de désengagement.

Cela suffit à modifier les actions suivantes de l'un et de l'autre — Marco recommence à remettre son travail à temps, Giulia relâche les contrôles — sans que leurs programmes respectifs (l'habitude d'évitement de Marco, le style de contrôle de Giulia) aient encore été touchés à la racine. Si le nouveau schéma se stabilise dans le temps, on aura en outre un véritable apprentissage : les programmes eux-mêmes, et non seulement les actions contingentes, se trouveront modifiés ; mais c'est là un résultat ultérieur, que le seul changement de perception ne garantit pas.

Une description purement systémique de cet épisode — un cycle de rétroaction négative entre les comportements de Marco et de Giulia — n'aurait pas été fausse, mais elle aurait laissé indistincts certains éléments qui se révèlent ici séparables : la différence entre l'exercice du pouvoir et la simple capacité d'agir ; la différence entre ce qu'un agent peut faire objectivement et ce qu'il perçoit pouvoir faire, qui est le point précis sur lequel la confrontation a agi ; le conflit entre deux programmes du même agent ou deux sous-agents, qui explique pourquoi Marco n'arrivait pas à faire ce que lui-même aurait voulu faire, sans avoir à invoquer autre chose que l'appareil déjà introduit ; et le constat, nullement évident, que la méta-interaction a fonctionné en modifiant des perceptions, non nécessairement les programmes sous-jacents. Ce sont des distinctions qu'un lexique fait de seuls entrées, sorties et rétroactions peut certes exprimer, mais qu'il ne nomme pas séparément ; et les avoir nommées séparément, ici, a produit les deux effets annoncés. Cela a indiqué à Giulia où agir : non sur la volonté de Marco, qui n'était pas en cause, mais sur la liste des actions qu'il croyait avoir à sa disposition — et c'est ce que la confrontation a fait, en y en ajoutant une. Et cela a dit aux deux à quoi s'attendre : la perception ayant changé et non le programme, l'amélioration est réelle mais fragile, et sous pression l'ancien comportement reviendra, si entre-temps rien d'autre n'est advenu. Sans cette distinction l'histoire aurait été classée comme résolue, et la rechute serait arrivée comme une surprise.

Conclusion

Pour comprendre quelque chose nous avons l'habitude de nous demander ce que c'est ou de quelle substance c'est fait : quel type de personne c'est, quel caractère elle a, comment est faite telle organisation. L'agentologie propose de changer les questions. Non pas ce qu'une chose est, mais qui agit, ce qu'il perçoit, ce qu'il veut, selon quelles logiques il prend ses décisions, quelles actions il croit avoir à sa disposition, etc.

Dire qu'une personne est paresseuse clôt la discussion : il ne reste qu'à le constater de nouveau et à le confirmer. Se demander en revanche ce que la personne voit lorsqu'elle se trouve devant une tâche, ce qu'elle craint, quelles options elle croit avoir, donne des réponses diversifiées, et chacune d'elles pourrait être différente si la situation changeait. Un jugement sur le caractère ne laisse rien à faire ; les questions agentologiques font apparaître des zones sur lesquelles on peut intervenir. Elles ne garantissent pas que l'intervention réussisse : elles disent où la tenter.

Les émotions, les rapports entre les personnes, les conflits internes. À première vue ces thèmes semblent n'avoir rien en commun, mais en chacun d'eux l'agentologie voit des agents qui interagissent selon des perceptions, des logiques et des objectifs, et elle propose les mêmes questions pour comprendre pourquoi advient ce qui advient.

Cet essai ne dit pas comment on doit traiter les autres, mais il invite à ne pas liquider une personne en disant qu'elle est faite d'une certaine manière, comme explication de son comportement. Cela ne signifie pas donner raison à n'importe qui, ni tout pardonner : celui qui a produit un dommage intentionnellement, alors qu'il pouvait agir autrement, en répond, et c'est ce qu'on appelle ici responsabilité. Cela signifie qu'on ne peut pas s'arrêter à la première explication qui vient à l'esprit, laquelle est d'ordinaire la plus commode et presque toujours inutile.

Je laisse au lecteur le soin de juger si l'appareil conceptuel ici proposé est utile. La prochaine fois qu'il verra quelqu'un faire une chose qui semble inexplicable, il pourra essayer de répondre aux questions du questionnaire agentologique. Si après cela il comprend quelque chose qu'il ne comprenait pas avant, le temps qu'il a passé à lire cet essai n'aura pas été perdu.

oOo