Rationalité limitée, calculabilité et complexité

Mikaël Cozic
mikael.cozic@ens.fr

(DEC, ENS Ulm)

Atelier "Théorie de la décision"
HEC, 2/III/06

introduction (1)

analyses computationnelles

- *calculabilité* : la fonction f est-elle calculable ?
- *complexité* : quelles ressources exige (le cas échéant) le calcul de la fonction f ?

rationalité limitée

3 composantes :

1. factuelle : les agents subissent des limitations cognitives
2. critique : les modèles de choix sont inadéquats compte tenu des limitations cognitives des agents
3. constructive : il faut construire des modèles de choix compatibles avec les limitations cognitives des agents

introduction (2)

analyses computationnelles et rationalité limitée

- les analyses computationnelles affirment être pertinentes pour la compréhension de la rationalité limitée (Kramer 1974, Richter & Wong 1999, Velupillai 2000)
- invocation des analyses computationnelles par les tenants de la rationalité limitée (Simon 1978)
- restrictions computationnelles en théorie des jeux répétés (Abreu et Rubinstein 1988, Rubinstein 1998, Neyman 1998)
- lacunes dans l'analyse méthodologique (Binmore 1987, Aumann 1997)

objet de l'exposé

- analyse et évaluation de l'apport des théories computationnelles à la rationalité limitée

introduction (3)

plan de l'exposé

1. analyses computationnelles et rationalité limitée
2. usage évaluatif des analyses computationnelles
3. usage constructif des analyses computationnelles

Section I

Analyse computationnelle et rationalité limitée

cadre épistémologique

cadre épistémologique

- modèle, modélisateur, champ d'application
  - modèle : structure formelle + interprétation générique
  - champ d'application : portion de réalité dont le modèle doit rendre compte, conformément à l'interprétation générique
- vertus descriptives vs. vertus pragmatiques :
  - vertus descriptives : capacités d'adéquation d'un modèle à son champ d'application
  - vertus pragmatiques : maniabilité du modèle par le modélisateur dans ses divers usages (prédictifs, explicatifs, etc.)

pertinence descriptive

analyses computationnelles des modèles

parce qu'il repose sur une structure formelle, tout modèle peut faire l'objet d'analyses computationnelles (*)

modèles physiques

- calculabilité : mécanique quantique (Pour-El et Richards, 1989)
- complexité : modèles d'Ising en physique statistique (Barahona, 1982 , Istrail 2000)

modèles de choix

- calculabilité : fonctions de choix du consommateur (Lewis, 1995 & 1992), équilibres généraux (Richter & Wong, 1999)
- complexité : modèles de choix de sous-ensembles (Fishburn & LaValle, 1996)

pertinence descriptive

- apport *commun* à l'ensemble des modèles : informations sur les vertus pragmatiques du modèle
- apport *spécifique* aux modèles de choix : informations sur les vertus descriptives du modèle = *hypothèse de pertinence descriptive*

décomposition de l'hypothèse de pertinence descriptive

- 2 composantes dans la justification de l'hypothèse de pertinence descriptive :

(1) connexion cognition-comportement

- raisonnons sur une fonction de choix $c(.)$ qui prend pour argument un ensemble d'actions réalisables X et pour valeur $c(X) \subseteq X$ (le comportement)
- une telle fonction de choix a une *contrepartie cognitive* chez les agents
- monotonie des adéquations cognitives et comportementales : plus un modèle de choix est plausible du point de vue cognitif, plus il est plausible du point de vue comportemental

(2) connexion computation-cognition

- lien entre les processus cognitifs et les analyses computationnelles
- les propriétés computationnelles comme indicateurs de la *difficulté* cognitive

limites

- *si* une fonction a une contrepartie cognitive, *alors* son analyse computationnelle a une pertinence descriptive
- *sinon*, les arguments qui précèdent ne garantissent pas la pertinence descriptive (*pace* Axtell, 2005)
- *exemple* : la non-calculabilité des équilibres généraux (Richter & Wong , 1999)

Section II
Usage évaluatif

1. théorie de la calculabilité
    1. exemple (non réalisabilité des fonctions de choix)
    2. discussion
2. théorie de la complexité (**NP**-difficulté du choix des sous-ensembles)
    1. exemple
    2. discussion

théorie de la calculabilité

- *explicandum* : une fonction $f : \mathbb{N}^p \to \mathbb{N}$ est calculable s'il existe un algorithme qui permet de donner en un nombre fini d'étapes la valeur de la fonction pour chacun de ses arguments
- *explicantes* : λ-définissabilité (Church, 1931), récursivité (Gödel, 1934), calculabilité par machine de Turing (Turing, 1936-7)
- équivalence extensionnelle entre les différents *explicantes* : f est λ-définissable ssi f est récursive ssi f est calculable par machine de Turing

machine de Turing

Définition

Une machine de Turing à m rubans infinis à droite est un quintuplet $M = \langle Q, \Sigma, q_0, \delta, F \rangle$:

(i) Σ *est un alphabet qui contient d (début de bande),* $|$ *et b (blanc) ;*

(ii) Q *est un ensemble fini d'états ;* q_0 *est l'état initial ;*

(iii) $\delta : Q \times \Sigma^m \rightarrow Q \times \Sigma^m \times \{-1, 0, +1\}$ *est la fonction de transition ;* -1 *indique que la tête de lecture se déplace d'une case vers la gauche,* 0 *qu'elle reste au même endroit et* $+1$ *qu'elle se déplace d'une case vers la droite ;*

(iv) F *est l'ensemble des états finaux ; il contient les états* q_y *(état acceptant),* q_n *(état refusant) et* q_a *(arrêt).*

Définition

Soit $f : \mathbb{N}^p \rightarrow \mathbb{N}$ *une fonction partielle et* M *une machine de Turing à* m *rubans avec* $m \geq p + 1$ *; on suppose un codage correct des entiers par les mots de* Σ *et on confondra par commodité les entiers et leurs codes.*

f **est calculée par** M *si, pour tout input* $(x_1, ..., x_p)$

(i) *quand* $f(x_1, ..., x_p)$ *n'est pas définie,* M *ne s'arrête pas*

(ii) *quand* $f(x_1, ..., x_p)$ *est définie,* M *s'arrête en un temps fini et sur le* $p + 1$*ème ruban figure le code de* $f(x_1, ..., x_p)$

On dit que f *est* **calculable par machine de Turing** *s'il existe une machine de Turing* M *qui calcule* f*.*

modèle de choix en situation de certitude

choix en situation de certitude

(M 1) l'agent peut choisir entre les éléments d'un ensemble d'opportunités ou d'actions réalisables A

(M 2) les préférences de l'agent sur les actions réalisables constituent un préordre total $\succeq \subseteq A \times A$ (total et transitif)

(M 3) l'agent choisit, s'il en existe au moins une, une des actions dont la conséquence est $\succeq$-maximale.

Définition

Soit A un ensemble d'actions réalisables et $\mathbb{F} \subseteq \wp(A)$; une **fonction de choix** *pour $\mathbb{F}$ est une fonction $c : \mathbb{F} \to \wp(A)$ t.q. $\forall X \in \mathbb{F},\ c(X) \subseteq X$. On appelle $(\mathbb{F}, c)$ une* **structure de choix**.

Définition

Soit $(\mathbb{F}, c)$ une structure de choix ;

(i) $(\mathbb{F}, c)$ *est* **rationalisée** *par une relation de préférence $\succeq$ si $\forall X \in \mathbb{F},\ c(X) = \{x \in X : \forall y x \succeq y\}$*

(ii) $(\mathbb{F}, c)$ *est* **rationalisable** *s'il existe une relation de préférence qui la rationalise.*

analyse récursive (1)

$\mathbb{R}$ (réels)	$\mathbb{R}_c$ (réels récursifs)

analyse récursive (2)

Définition

Une suite de rationnels $\{r_x\}$ *est récursive s'il existe trois fonctions récursives* $sign, den, num$ *telles que pour tout* $x \in \mathbb{N}$,
$r_x = (-1)^{sign(x)} num(x)/den(x)$

Définition (réels récursifs)

Un nombre réel α *est un* **réel récursif** *s'il existe une suite de rationnels récursive* $\{r_x\}$ *et une fonction récursive* f *telle que pour tout* n
si $x \geq f(n)$*, alors* $|r_x - \alpha| \leq 2^{-n}$

$\mathbb{R}$ (réels)	$\mathbb{R}_c$ (réels récursifs)
$A \subseteq \mathbb{R}^n$ (ensemble des actions réalisables) $\mathbb{F} = \{X \subseteq A\}$ (sous-ensembles d' actions réalisables)	$R(A) \subseteq M(\mathbb{R}^n)$ (ensemble récursif des actions réalisables) $\mathbb{F}_R = \{X : X \subseteq R(A) \wedge$ X récursif$\}$ (sous-ensembles récursifs d'actions réalisables)

Définition

Un **espace métrique récursif** *est une paire* $(T, \sim_T)$ *(où* $T \subseteq \mathbb{N}$ *et* $\sim_T$ *est une relation d'équivalence sur* T*) et un opérateur binaire récursif* $D : T \to \mathbb{R}$ *tel que*

(i) $\forall a, b \in T[(D(a, b) = 0) \Leftrightarrow (a \sim_T b)]$

(ii) $\forall a, b \in T[D(a, b) = D(b, a)]$

(iii) $\forall a, b, c \in T[D(a, b) \leq D(a, b) + D(b, c)]$

O

n peut alors former les espaces métriques récursifs $M(\mathbb{R}) = ((\mathbb{N}(R), \sim_{\mathbb{N}(R)}), D_{\mathbb{R}})$ et $M(\mathbb{R}^n)$ à partir de l'ensemble $\mathbb{N}(R)$ des nombres de Gödel des (fonctions f associées aux) réels récursifs.

Soit $A \subseteq \mathbb{R}_+^n$ compact et convexe ; on définit $R(A)$, sous-ensemble de l'espace métrique récursif $M(\mathbb{R}_+^n)$

Définition

$\mathbb{F}_R$ est l'ensemble des sous-ensembles récursifs de $R(A) = \{X \in \wp(R(A))$ et X est récursif $\}$; la paire $(R(A), \mathbb{F}_R)$ est l'ensemble des **sous-ensembles récursifs d'actions réalisables**. *On note $\{\mathbb{F}_{Rj}\}_{j \in \mathbb{N}}$ une sous-famille de $\mathbb{F}_R$ effectivement énumérable.*

$\mathbb{R}$ (réels)	$\mathbb{R}_c$ (réels récursifs)
$A \subseteq \mathbb{R}^n$ (ensemble des actions réalisables) $\mathbb{F} = \{X \subseteq A\}$ (sous-ensembles d'actions réalisables)	$R(A) \subseteq M(\mathbb{R}^n)$ (ensemble récursif des actions réalisables) $\mathbb{F}_R = \{X : X \subseteq R(A) \wedge$ X récursif$\}$ (sous-ensembles récursifs d'actions réalisables)
$c : \mathbb{F} \longrightarrow \wp(A)$	$c : \mathbb{F}_R \longrightarrow \wp(R(A))$

Définition

Une fonction de choix c sur $(R(A), \mathbb{F}_R)$ est **récursivement rationalisable** *s'il existe*

(i) *une relation* $\succeq: R(A) \times R(A) \to \{1, 0\}$

(ii) *une fonction* $f : R(A) \to \mathbb{N}$ *récursive partielle telle que* $\forall a, b \in R(A)[(a \succeq b) = 1 \to f(a) \geq f(b)]$ *et* $\forall X \in \mathbb{F}_R$, $c(X) = [a : \forall b \in X(f(a) \geq f(b))]$.

Définition

Pour un domaine $\{\mathbb{F}_{Rj}\}_{j\in\mathbb{N}} \subseteq \mathbb{F}_R$ *et un co-domaine* $\{c(\mathbb{F}_{Rj})\}_{j\in\mathbb{N}}$*, le* **graphe** *de c est la collection des paires* $(\mathbb{F}_{Rj}, c(\mathbb{F}_{Rj}))$*. Le graphe de c a un* **domaine complet** *si pour un certain* $K \in \mathbb{N}$ *et pour toutes paires* $i \neq j > K$, $\mathbb{F}_{Ri} \triangle \mathbb{F}_{Rj} \neq \emptyset$.

Définition

Une fonction de choix sur $(R(A), \mathbb{F}_R)$ *qui est récursivement rationalisable est* **récursivement réalisable** *ssi pour tout domaine complet* $\{\mathbb{F}_{R_j(j\in\mathbb{N})} \subseteq \mathbb{F}_R\}$*, le graphe de C est un ensemble récursif de l'espace* $\wp(M(\mathbb{R}^n)) \times \wp(M(\mathbb{R}^n))$.

résultat d'impossibilité

Théorème (Lewis, 1985)

Soit c est une fonction de choix récursivement rationalisable non-triviale sur $(R(A), \mathbb{F}_R)$*, alors C n'est pas récursivement réalisable et* $\{\mathbb{F}_{Rj}\}$ *un domaine complet. Le graphe de c n'est pas récursivement réalisable.*

discussion (1)

situation générique

- modèle de choix M selon lequel à toute situation de choix S, l'agent associe une action déterminée par la solution $SOL_M(.)$ du modèle : $M : S \rightarrow SOL_M(S)$
- typiquement, $SOL_M(.)$ consiste à sélectionner une action optimale relativement à une relation de préférence ou à maximiser une fonction d'utilité
- $SOL(.)$ non calculable par MT

question

Quel est l'impact du résultat de non-calculabilité sur l'évaluation de M ?

discussion (2)

(P1) charge critique

non calculabilité par MT

$\Downarrow$

impossibilité computationnelle

$\Downarrow$

improbabilité cognitive

$\Downarrow$

improbabilité comportementale

discussion (3)

(P2) charge de la preuve

- option 1 : attaquer l'argument de la non-calculabilité par MT à l'improbabilité comportementale
- option 2 : faire valoir que l'objectif du modèle M est de fournir une approximation satisfaisante du comportement
- inversion de la charge de la preuve

théorie de la complexité

- idéalisation de la notion de calculabilité / calculabilité *en pratique* ou *faisabilité*
- la théorie de la complexité construit des notions plus proches de la notion de calculabilité en pratique
- mesure des ressources en temps ou en espace
- **P** et **NP**

résultat de *NP*-difficulté

modèle de choix de sous-ensemble

- ensemble fini d'opportunités A
- fonction d'utilité linéaire $u(X) = \sum_{a \in X} u(a)$
- solution $Sol = \arg\max_{X \subseteq A : p(X) < w} u(X)$

Proposition (Fishburn & LaValle 1996)

Sol est **NP***-difficile.*

discussion

Test computationnel de M :

Etape 1 : on détermine une classe $\mathfrak{F}_P$ de fonctions cognitivement plausibles.

Etape 2 : M est soumis à un *test computationnel relativement à* $\mathfrak{F}_P$: M passe le test si les fonctions qui constituent M (ou que constitue M) appartiennent à $\mathfrak{F}_P$.

Etape 3 : si ce n'est pas le cas, il faut ou bien trouver des raisons spéciales pour supposer que le modèle est empiriquement adéquat dans le champ d'application visé, ou bien réviser le modèle.

Section III
Usage constructif

dilemme du prisonnier

		Colin	
		C	D
Louise	C	(3,3)	(0,4)
	D	(4,0)	(1,1)

jeux répétés à horizon fini

Définition

Soit $G = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ un jeu sous forme normale ;

(i) *une* **histoire à l'étape** *k de la répétition de G, notée H^k est un élément de A^{k-1}; on pose $A^0 = \emptyset$.*

(ii) *l'ensemble des* **stratégies pures du joueur** *i* **dans le jeu** *G^t est l'ensemble des fontions $\sigma_i : \bigcup_{k=1}^{t} H^k \rightarrow A_i$*

(iii) *chaque profil $\sigma = (\sigma_1, ..., \sigma_N)$ de stratégies pures pour G^t induit une partie à t étapes $\omega(\sigma) = (\omega_1(\sigma), ..., \omega_t(\sigma))$ définie comme suit:*

- $\omega_1(\sigma) = \sigma(\emptyset) = (\sigma_1(\emptyset), ..., \sigma_N(\emptyset))$ *et*
- $\omega_k(\sigma) = \sigma(\omega_1(\sigma), ..., \omega_{k-1}(\sigma)) = (\sigma_1(\omega_1(\sigma), ..., \omega_{k-1}(\sigma)), ..., \sigma_N(\omega_1(\sigma), ..., \omega_{k-1}(\sigma))$

Définition (suite)

(iv) *soit σ un profil de stratégies pures pour G^t; on définit alors les paiements du joueur i quand σ est joué comme*

$$u_i^t(\sigma) = \frac{\left(\sum_{k=1}^{\infty} u_i(\omega_k(\sigma))\right)}{t} \tag{1}$$

Définition

Soit $G = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ un jeu sous forme normale ;

(i) *une* **machine pour le joueur** *i est un 4-uplet $m_i =< Q, q_0, f, t >$ où*

- *Q est l'ensemble des états*
- *$q_0 \in Q$ est l'état initial*
- *$f : Q \longrightarrow A_i$ est la fonction de sortie*
- *$\tau : Q \times A_{-i} \longrightarrow Q$ est la fonction de transition*

On note M_i l'ensemble des machines pour le joueur i.

(ii) *une* **machine aléatoire pour le joueur** *i est une distribution de probabilités sur l'ensemble des machines pour le joueur i soit un élément de $\Delta(M_i)$.*

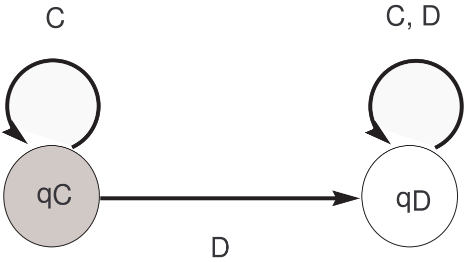

Figure: Un automate pour "Oeil pour Oeil"

Définition

Soit $G = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$ *un jeu sous forme normale ;*

(i) $r_i : \mathbb{N} \to \mathbb{N}$ *est une fonction de restriction qui assigne au joueur* i *la complexité maximale* $r_i(t)$ *des stratégies qu'il peut choisir dans le jeu répété* G^t

(ii) $G^t(r_1, \ldots, r_I)$ *le jeu répété dont (a) les stratégies du joueur* i *sont les stratégies* σ_i *telles que* $c(\sigma_i) \leq r_i(t)$ *et (b) les paiements sont la restriction des paiements de* G^t *aux stratégies restantes*

(iii) *si* G *est un jeu, on note* $Eq(G)$ *l'ensemble des équilibres de Nash de* G *et* $pEq(G)$ *l'ensemble des profils de paiements des équilibres de* G

Théorème (Papadimitriou & Yannakakis, 1994)

Soit G^t un dilemme du prisonnier répété t fois ; pour tout $\epsilon > 0$, si $r_1(t)$ ou $r_2(t) \leq 2^{c_\epsilon t}$ avec $c_\epsilon = \frac{\epsilon}{12(1+\epsilon)}$, alors pour un t assez grand, il existe dans $G^t(r_1, r_2)$ un équilibre de Nash de paiement moyen au moins égal à $3 - \epsilon$

Théorème (Neyman, 1998)

Soit un jeu $G = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$*; pour tout* $\varepsilon > 0$ *assez grand il existe deux entiers positifs* t_0 *et* r_0 *tels que (a) si* $t \geq t_0$*, (b)* w *est un profil de paiements strictement individuellement rationnel et (c)* $r_0 \leq \min(r_1, r_2) \leq \exp(\varepsilon^3.t)$*, alors il existe un équilibre de Nash dont le profil de paiements* $y \in pEq(G^t(r_1, r_2))$ *est tel que pour tout joueur* i*,* $|y_i - w_i| < \varepsilon$*.*

discussion (1)

stratégies admissibles vs. accessibles

- stratégies admissibles : stratégies retenues par le modèle
- stratégies accessibles : stratégies auxquelles les agents ont effectivement accès

attente faible vs. forte

- attente faible : l'objectif de la restriction est d'introduire une variable cognitive
- attente forte : l'objectif de la restriction est d'approcher par les stratégies admissibles les stratégies accessibles

discussion (2)

2 exigences

1. exigence de support empirique
2. exigence de systématicité

Théorème (Papadimitriou, 1992)

Soit G^t un dilemme du prisonnier répété à t périodes ; étant donné un automate fini m_1 et un entier k, le problème qui consiste à trouver un automate m_2 de taille inférieur à k et de paiement optimal est **NP***-difficile.*

conclusion (1)

asymétrie

- asymétrie entre usage évaluatif et usage constructif. Pourquoi ?
- facteurs contingents : calculabilité ou calculabilité en temps polynomial vs. contraintes de cardinalité sur les états internes d'automates finis
- facteurs structurels : l'amendement opéré par les restrictions computationnelles est extrêmement local
  $\Rightarrow$ question : comment amender le critère d'optimisation ?

conclusion (2)

problème de l'amendement du critère de choix

- considérons de nouveau un modèle M bouclé par une solution (critère d'optimisation) $SOL_M(.)$
- supposons que la classe des fonctions cognitivement plausibles soit $\mathbb{F}_P$ et que $SOL_M \notin \mathbb{F}_P$
- par quoi remplacer SOL_M ?

conclusion(3)

prolongements de l'usage évaluatif

1. rapport complexité (notions computationnelles) / difficulté (capacités cognitives)
2. test computationnel des modèles de rationalité limitée (ex : le *satisficing*)

références (1)

D. Abreu et A. Rubinstein, "The Structure of Nash Equilibrium in Repeated Games with Finite Automata, *Econometrica*, 1988, 56, 6, pp. 1259-1281

R. Aumann, "Rationality and Bounded Rationality", *Games and Economic Behavior*, 1997, 21, pp. 2-14

G. Kramer, "An Impossibility Result Concerning the Theory of Decision Making", Yale University, 1974, Cowles Foundation Reprint, 274

A. A. Lewis, "On Effectively Computable Realizations of Choice Functions", *Mathematical Social Sciences*, 1985, 10, pp. 43-80

A. Neyman, "Finitely Repeated Games with Finite Automata", *Mathematics of Operation Research*, 1998, 23, 3, pp. 513-552

références (2)

C. Papadimitriou et M. Yannakakis, "On Bounded Rationality and Computational Complexity", STOC 94, ACM

M.R. Pour-El et J.I. Richards, *Computability and Analysis in Physics*, Springer-Verlag, 1989

A. Rubinstein, *Modeling Bounded Rationality*, MIT Press, 1998

K. Velupillai, *Computable Economics*, Oxford UP, 2000