Part II
Principe de réfutation et processus de découverte

3. Introduction

Nous allons examiner dans les parties suivantes, certaines conséquences des thèses popperiennes, sur les processus de découverte au sein d'une communauté d'agents. En remplaçant vérité par corroboration dans la recherche scientifique, la théorie de Popper fait naître un nouveau problème dans le déroulement de la recherche scientifique. En effet, si la recherche a pu paraître orientée vers un but précis, l'obtention d'énoncés vrais, ce but est placée hors d'atteinte par l'épistémologie poppérienne et du même coup, il n'y a plus aucun indicateur susceptible de déterminer l'arrêt d'une recherche de manière certaine. Une recherche s'arrête quand la théorie est suffisamment corroborée. Mais par rapport à quoi est évalué ce suffisamment? La réponse que semble inspirer Popper est que le chercheur ne peut évaluer l'opportunité de l'abandon d'un processus de recherche qu'en fonction de considérations d'utilité qui prendront en compte notamment des facteurs sociaux et les intérêts personnels. Il serait alors intéressant de voir comment sont liés le développement des connaissances d'une communauté et les comportements individuels des agents. Quel est l'influence, à travers les compromis exploration-exploitation que font les agents, des paramètres sociaux sur le développement de la connaissance?

Selon la vision poppérienne, l'organisation de l'activité scientifique d'une communauté suit un schéma relativement simple. Chaque scientifique établit des théories qu'il teste pas à pas jusqu'à ce qu'elle soit réfutée ou qu'elle soit suffisamment corroborée pour mériter d'être proposée à la communauté. Ensuite, si elle est acceptée, c'est à dire si aucun autre membre de la communauté n'est capable de la réfuter et si elle présente un interêt réel. cette théorie figure au nombre de l'ensemble des théories acceptées par la communauté tant qu'elle n'est pas réfutée (nous noterons cet ensemble CP pour connaissances publiques). Il est alors tentant de modéliser ces mécanismes afin d'en voir les conséquences sur l'évolution des connaissances admises. C'est ce que nous nous proposons de faire.

Nous allons pour cela nous placer dans le cadre d'une société multi-agents gouvernée par les règles de la théorie des jeux. Les agents représenteront les scientifiques. Ceux-ci seront rétribués en fonction de leurs actions. Les gains et les pertes seront des variables réelles. Pour guider leur actions, les agents chercheront à optimiser leurs gains. Cette règle ne fait

aucune assertion sur la psychologie des agents, comme on pourrait le penser, au premier abord . En effet, les gains peuvent faire jouer des aspects non physiquement quantifiables, tels que la réputation ou des élément culturels. De ce point de vue, la seule chose qu'implique le principe d'optimisation des gains, est que les individus ont des préférences et sont capables de hiérarchiser les conséquences de leurs actions en fonction de leur état cognitif du moment. Il serait d'ailleurs possible de considérer des espaces de gains plus compliqués que l'axe des réels, tel un treillis, sur lequel l'ordre n'est pas total. Ceci permettrait d'avoir des éléments maximaux non comparables, ce qui serait plus proche d'une situation réelle. Les limites de cette règle sont donc pour la plupart théoriques et computationnelles.

L'idée est alors de considérer un espace abstrait de théories duquel les agents tirent des éléments. La validité de chaque théorie ainsi choisie peut être mise à l'épreuve par les agents qui effectuent des tests, afin de confronter les prédictions de la théorie aux données empiriques. A tout moment, un agent peut décider de proposer à la communauté la théorie qui fait l'objet de ses recherches. Suivant les connaissances de la communauté, celle-ci choisit de reconnaître la théorie proposée, en l'incorporant au connaissances communes (CP), et de récompenser l'agent qui propose de G, ou bien de la rejeter et de pénaliser l'agent de $-R$. Les agents ont une durée de vie finie, à chaque pas de temps, ils ont la probabilité $1-\alpha$ de cesser toute activité. Si du vivant d'un agent, un autre agent trouve un contre exemple à une théorie qu'il a proposée, ce dernier gagne R et le premier perd R. Nous étudierons en particulier, le compromis exploration-exploitation auquel est confronté l'agent, qui doit choisir entre explorer les théories possibles et les manières de les tester, et exploiter les théories qui auront résistées suffisamment longtemps à l'épreuve des tests. Ceci nous amènera à déterminer la forme d'un comportement optimal.

4. Présentation du modèle

Considérons un espace de théorie Φ et un ensemble d'énoncés $\mathcal{E}$ de base appartenant à la réunion des champ d'application des théories considérées. A chaque couple $(\varphi, e) \in \Phi \otimes \mathcal{E}$ nous pouvons associer une valeur 1 ou 0 selon que la théorie φ est ou n'est pas compatible avec l'énoncé de base e. L'ensemble $\mathcal{E}_\varphi$ des énoncés compatibles avec une théorie φ sera alors l'ensemble des énoncés de base sur lesquels la théorie vaut 1 . L'ensemble $\mathcal{E}_{\overline{\varphi}}$ des falsificateurs virtuels de la théorie sera l'ensemble des énoncés sur lesquels la théorie vaut 0. L'ensemble $\mathcal{E}_{\overline{\varphi}}^{réel}$ des contre-exemples à la théorie sera l'ensemble des énoncés qui décrivent une partie du monde réel et sur lesquels la théorie vaut 0. Il faut remarquer qu'aucune assertion n'est faite quant à la forme des théories envisagées. Celles-ci peuvent être des énoncés linguistiques, des équations mathématiques, des schémas, ou quoi que ce soit qui réponde à l'exigence de Popper à savoir, que l'ensemble des falsificateurs virtuels de la théorie soit non vide. Nous allons dans cette partie rester le plus général possible. Nous spécifierons quelques paramètres dans la partie suivante. Nous nous placerons dans le cadre

du calcul des probabilités.

4.1. Préliminaires mathématiques

Avant de commencer une formalisation, nous allons exposer quelques concepts de base du calcul des probabilités concernant l'acquisition d'information.

Les tribus. Suivant les conceptions de Popper, un scientifique teste les théories pas à pas. A partir des informations accumulées, il catégorise l'espace des théories. C'est à partir de cette catégorisation qu'il évalue leur valeur. Nous allons donc voir comment il est possible de traduire mathématiquement ce processus de catégorisation. Pour illustrer ceci, prenons l'exemple de quelqu'un qui reçoit un cadeau ω et essaie d'en deviner le contenu avant de le déballer. L'emballage se présente sous la forme d'une petite boîte rectangulaire. Les couleurs préférées de cette personne étant le bleu et le jaune, plusieurs possibilités s'offrent à lui. Il peut s'agir d'un stylo, jaune ou bleu, ou une montre, jaune ou bleue. Pour déterminer le contenu, il peut poser comme question 1_ "est-ce une montre?", 2_"est-ce un stylo?", 3_ "est-ce jaune?", 4_"est-ce bleu?". L'espace des évènements est ici constitué des contenus possibles de l'emballage. Il comporte donc quatre éléments. Le contenu réel ω est a priori aléatoire. Le comportement de la personne est déterminé par l'ordre dans lequel elle va poser les deux questions qui permettront de déterminer le contenu. Supposons qu'elle décide de poser ses questions dans l'ordre 1,3. Nous avons alors les possibilités suivante:

1. ω est une montre. A la première question, il reste deux possibilités, montre jaune et montre bleue . A la deuxième question toute l'information sur ω est disponible.

2. ω est un stylo. A la première question, il reste deux possibilités, stylo jaune et stylo bleu. A la deuxième question toute l'information sur ω est disponible.

Nous voyons donc qu'à chaque étape, l'état cognitif de la personne est caractérisé par une collection de catégories disponibles: (montre jaune ou bleue) , (stylo jaune ou bleu), et toutes les catégories composées à la première étape; montre jaune, montre bleue, stylo jaune, stylo bleu et toutes les catégories composées à la seconde étape. A chaque étape, l'information disponible sur l'objet est donnée par la plus petite catégorie ou classe d'équivalence qui contient l'objet. Par exemple, si ω est une montre, l'information à la première étape est (montre jaune ou bleue). L'acquisition d'information s'interprète donc comme une catégorisation de plus en plus fine sur l'espace des évènements.

De manière plus abstraite, à un espace d'évènement Φ et à un mécanisme d'acquisition d'information, nous pouvons associer une famille de catégorisations sur Φ, que nous appellerons des tribus, et qui seront indexées par le temps. Nous les noterons $\mathcal{B}^t$. Les $\mathcal{B}^t$ seront des collections emboitées de plus en plus fines, d'ensembles engendrés par des classes d'équivalence sur Φ. Nous avons ainsi pour $t < t'$, $\mathcal{B}^t \subset \mathcal{B}^{t'}$. A un instant donné t, le

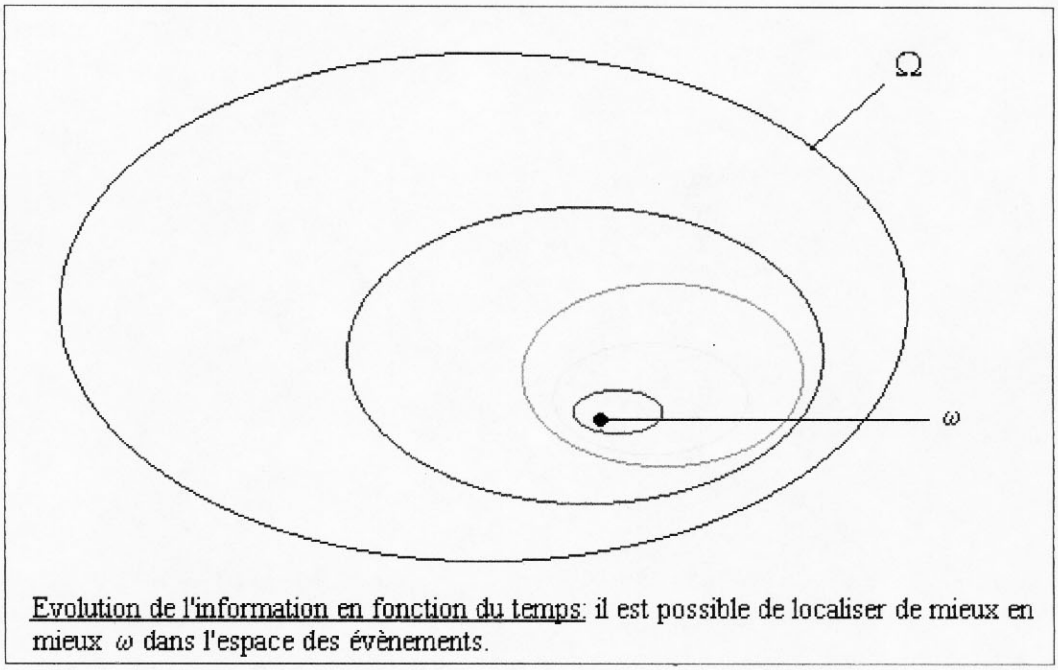


Figure 3:

maximum d'information disponible à propos d'un événement aléatoire ω est porté par le plus petit élément de la tribu $\mathcal{B}^t$ qui le contient. Nous pouvons nous représenter les choses graphiquement à l'aide de la figure 3. Un espace Φ munit d'un processus d'acquisition d'information engendrant les familles de classes d'équivalence $\mathcal{B}^t$ sera noté $(\Phi, \mathcal{B}^t)$.

Les temps d'arrêt. Au sein d'un processus de recherche, un agent doit déterminer le moment le plus favorable pour interrompre la recherche et choisir une nouvelle théorie. Cette décision sera prise en fonction de l'information accumulée qui se trouve être ici, le nombre de tests corroborant réalisés, ainsi que le nombre total de tests effectués. Une stratégie consistera donc à se donner pour chaque type de processus de recherche un temps d'arrêt. Ce sont ces temps d'arrêt qui détermineront les actions des agents. Leur étude sera donc essentielle pour comprendre la dynamique de la société.

La notion de temps d'arrêt en probabilité étant assez subtile, nous allons l'expliquer plus en détail. Un temps d'arrêt ν sur un espace $(\Phi, \mathcal{B}^t)$ est une variable aléatoire sur Φ à valeurs dans $\mathbb{N}$. . Un agent qui tire un élément $\omega \in \Phi$ acquiert petit à petit des informations sur ω qui lui permettent de se faire une idée quant à sa nature (dans ce qui suit, l'agent tire une formule et une suite d'expériences et recherche à évaluer le type de la formule). Les informations disponibles à chaque instant sont représentées mathématiquement par les tribus qui donnent toutes les catégorisations possibles que peut effectuer un agent sur l'espace des évènements. Le temps d'arrêt donne pour tout évènement ω, le moment $n = \nu(\omega)$ où l'agent arrêtera d'accumuler de l'information. A l'instant n, l'agent dipose donc de classe d'équivalences

sur Φ, l'information la plus précise qu'il pourra avoir sur ω sera alors la plus petite classe d'équivalence contenant ω, que nous noterons $C^n(\omega)$. L'agent sait au temps n que ω est l'un des élément de $C^n(\omega)$ et que ω possède de manière certaine les propriétés communes aux éléments de cette classe. Ce qu'il faut retenir de tout ceci, c'est qu'à l'instant n, l'agent ne peut pas discerner ω des autres éléments de $C^n(\omega)$. En quelque sorte, $C^n(\omega)$ est pour lui un point à part entière dans l'espace des énènements. Mathématiquement, une variable aléatoire est un temps d'arrêt si elle respecte les classes d'équivalences définies par les différentes tribus. Tout ceci impose une condition sur les temps d'arrêt : si pour $\omega \in \Phi$, $\nu(\omega) = n$, alors ν est égal à n pour tout autre élément de $C^n(\omega)$. Autrement dit, puisque les éléments de $C^n(\omega)$ ne sont pas discernables à l'intant n, l'agent est obligé de les traiter de la même manière. Nous obtenons ainsi l'expression mathématique des temps d'arrêt: Soit ν une v.a. sur $(\Phi, \mathcal{B}^t)$, ν est un temps d'arrêt si et seulement si:

$$\forall n \in \mathbb{N}, \{\omega \mid \nu(\omega) = n\} \in \mathcal{B}^n$$

Cette contrainte sur les temps d'arrêt fait que ceux-ci ne peuvent pas avoir n'importe quelle forme. Les agents cherchant les temps d'arrêt qui optimisent leurs gains, il nous faudra déterminer la manière dont les agents évaluent leur situation à un moment donné, c'est à dire, les espérances de gain associées aux différentes actions qui s'offrent à eux.

4.2. Notations

Espace des théories.

- Φ ensemble des théories.
- $\mathcal{E}$, ensemble des expériences possibles. $\mathcal{E}^{réel}$ désignera le sous ensemble de $\mathcal{E}$ des énoncés de base décrivant le monde réel.
- $\Pi = \{\pi \mid \pi : \mathbb{N} \to \mathcal{E}\}$, ensemble des suites de tests possibles.
- Pour $\varphi \in \Phi$ et $e \in \mathcal{E}$, $\varphi(e) \in \{0, 1\}$ sera la compatibilité de φ avec l'expérience e.
- $\Omega = \Phi \otimes \mathcal{E}$ ensemble des tirages possibles d'évènements aléatoires. Un élément (φ, π) de Ω sera noté ω.

Les théories.

- Pour une théorie φ, nous noterons $\overline{\varphi}$ la théorie opposée, celle qui prend pour phénomènes compatibles les falsificateurs virtuels de φ
- A chaque théorie, nous pouvons associer l'ensemble des énoncés de base compatibles avec celle-ci: $\mathcal{E}_\varphi$, et les falsificateurs virtuels: $\mathcal{E}_{\overline{\varphi}}$.

- A l'ensemble $\mathcal{E}_\varphi \cap \mathcal{E}^{réel}$, que nous noterons $\mathcal{E}_\varphi^{réel}$, nous associerons un réel $\theta_\varphi \in [0,1]$ représenterant la probabilité pour qu'une expérience du monde réel vienne réfuter φ. Nous avons donc $\theta_\varphi = \frac{\left|\mathcal{E}_\varphi^{réel}\right|}{\left|\mathcal{E}^{réel}\right|}$. Nous considérerons que l'ensemble $\{\theta_\varphi\}_{\varphi\in\Phi}$ est un ensemble fini. Ceci n'est pas très contraigant puisque nous pouvons considérer un nombre de types aussi grand que l'on veut et varier leur concentration autour des valeurs qui nous intéressent. D'ailleurs, si la précision nécéssairement finie des mesures que nous effectuons dans le monde réel discrétise en fait, le monde phénoménal.

  Les θ_φ étant en nombre fini, nous pouvons les ordonner en une suite croissante $(\theta_k)_{k\in I}$, et ensuite classer les théories en types $T_1, ..., T_q$. La classe des formules toujours vraies dans le monde réel sera notée T_∞.

la communauté.

- τ l'âge de la communauté

- A la population des agents, $|A| = N$. Nous poserons $r = \frac{N}{|\Phi|}$, c'est un indice de la richesse du champ de recherche.

- CP, connaissances partagées de la communauté : c'est l'ensemble des théories qui ont fait l'objet d'une publication. Cet ensemble se subdivise en deux sous ensembles, l'ensemble F des théories qui ont été prouvées fausses et l'ensemble V des théories acceptées par la société et donc considérées comme "vraies".

Dans tout ce qui suit, nous noterons pour une variable aléatoire v, $E^t(v)$ l'espérance de v, conditionnellement à l'information disponible au temps t et $E(v)$ l'espérance de cette variable en l'absence de toute information.

4.3. Organisation de la communauté

Horizon d'un agent. Chaque agent ayant la probabilité $1-\alpha$ de sortir du jeu à chaque pas de temps, la probabilité qu'il vive pendant t pas de temps à un instant donné est $p_t = (1-\alpha).\alpha^t$. La durée de vie moyenne d'un agent est donc $\langle T\rangle = (1-\alpha)\sum_{t=0}^{\infty} t.\alpha^t$. $\langle T\rangle = \left[\frac{\alpha}{1-\alpha}\right]$

Comportement des agents. Chaque agent est caractérisé par une suite infinie $H \in (\Phi\otimes\Pi)^{\mathbb{N}}$. C'est l'histoire de l'agent. H représente dans l'ordre, toutes les théories que l'agent sera susceptible d'étudier au cours de sa vie et pour chaque théorie, H donne dans l'ordre, tous les tests que l'agent sera susceptible de lui faire subir. Il ne faut pas être trompé par la formalisation, croire que tout est déjà écrit d'avance et que les agents subissent leur destin. Une des caractéristiques des probabilités étant de considérer toutes les suites d'évènements possibles, il est équivalent de dire que tout est écrit d'avance ou que les choses arrivent au fur et à mesure. L'activité d'un agent consiste alors à découvrir son histoire en

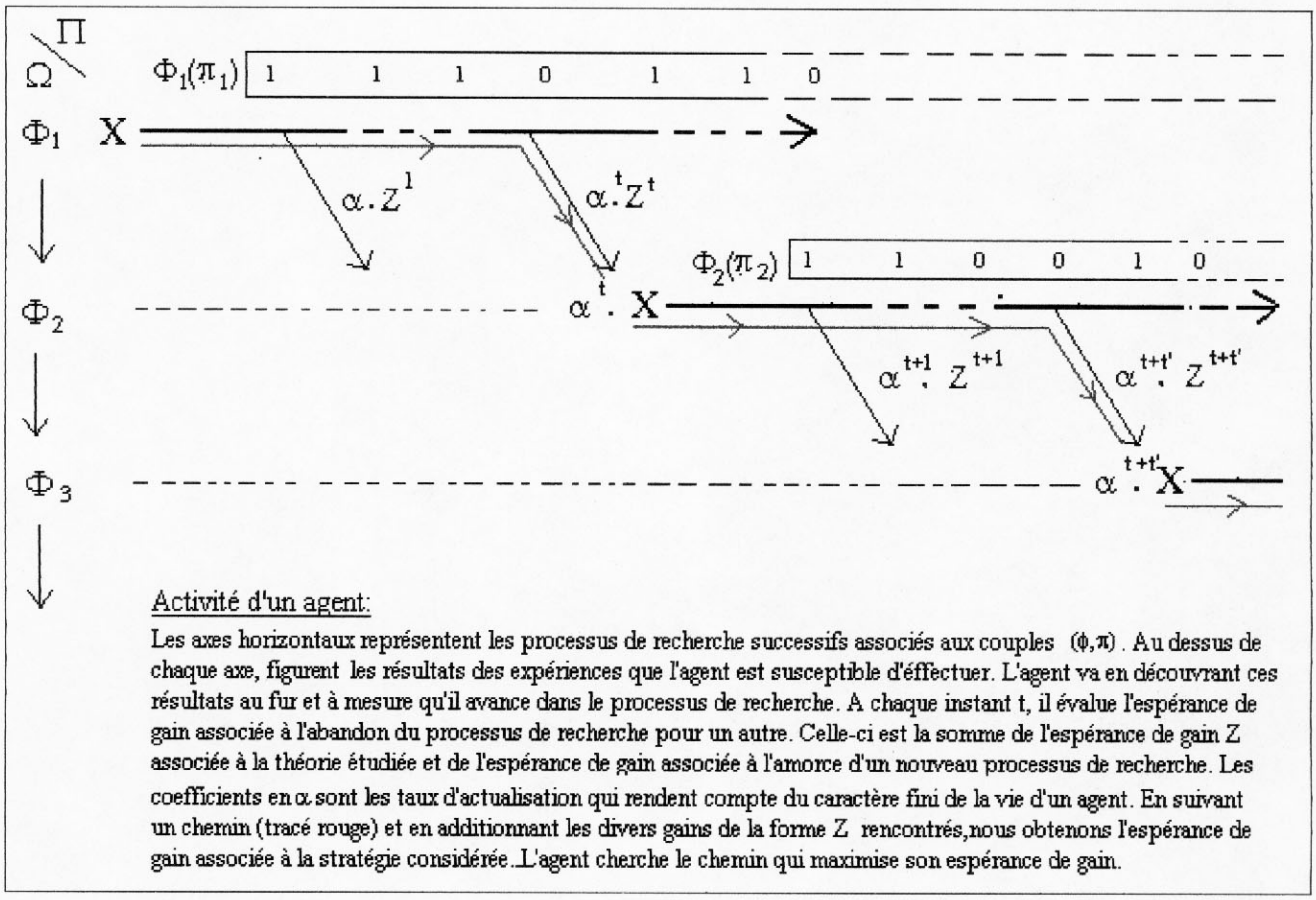


Figure 4:

essayant d'en tirer le plus de profit, c'est à dire qu'il découvre une à une les théories qui figurent dans H, théories qu'il teste pas à pas jusqu'à une profondeur qu'il détermine en fonction de ses motivations. La vie d'un agent est donc constituée d'un enchaînement de processus de recherche de différentes natures. La nature du processus de recherche dépend du statut de la théorie au moment où elle est tirée, c'est à dire, si elle est entièrement nouvelle ou si elle figure déjà dans CP. Chaque type de processus de recherche est caractérisé par une espérance de gain initiale qui évalue les conséquences d'une recherche guidée par ce processus. Quand il n'a pas encore choisi de direction de recherche en l'absence de toute information, l'agent dispose d'une espérance de gain initiale $G.X$. $G.X$ représente ce que peut attendre l'agent de son avenir d'après les principes qu'il s'est donné pour mener ses recherches. Nous verrons plus loin pourquoi il est naturel de mettre G en facteur. A tout moment t, au sein d'un processus, plusieurs actions sont possibles et les agents disposent d'évaluations Z^t des espérances de gains relatives à l'exploitation des informations acquises lors du processus de recherche en cours. Ceci est illustré par la figure 4.

Diverses actions possibles. Lors de l'évolution de la communauté, trois sous domaines de l'espace des formules évoluent paralèllement. Nous pouvons en effet distinguer:

- L'ensembles $\overline{CP}$ des formules qui n'ont pas encore été publiées, en proportion $(1-\sigma)$ dans Φ, caractérisé par la distribution $(\overline{\mu}_k)_k$ sur les types. Un processus de recherche dans ce domaine est caractérisé par une espérance de gain initiale $\overline{X}_0$.
- L'ensemble V des formules qui sont reconnues parcequ'elle ont été proposées et non réfutées, en proportion $\sigma\theta$ dans Φ, caractérisé par la distribution $(\mu_k)_k$. Un processus de recherche dans ce domaine est caractérisé par une espérance de gain X_0.
- L'ensemble F des formules qui ont été réfutées, en proportion $\sigma(1-\theta)$, caractérisé par la distribution $(\mu'_k)_k$.

Les agents entament des processus se déroulant sur l'échelle de temps t comme suit:

- A chaque pas de temps, un agent à la probabilité $1-\alpha$ de sortir du jeu.
- au temps $t=0$, la Nature tire aléatoirement un couple $\omega=(\varphi,\pi)\in(\Phi,\Pi)$, l'agent est informé du fait que $\varphi\in CK$ si c'est le cas. $\omega\in H$, les point de vue selon lesquels H est tirée au début de la vie une fois pour toute ou tirée au fur et à mesure de la vie de l'agent sont équivalent. Deux possibilités s'offrent alors:
  1. $\varphi\in\overline{CP}$, il peut décider de retirer une nouvelle formule ou bien entamer un processus de recherche. Si tel est le cas, il effectue des expériences jusqu'au temps d'arrêt $\overline{\nu}(\omega)$. A une expérimentation effectuée au temps t, l'agent récolte l'information supplémentaire $\varphi(\pi(t))\in\{0,1\}$, qui indique si la théorie est compatible avec les résultats de l'expérience réalisée. Lorsqu'il décide d'arrêter de chercher il a le choix entre deux possibilités:
     (a) proposer φ et gagner G au risque de proposer une formule fausse qui pourrait lui coûter R si quelqu'un propose un contre-exemple ultérieurement.
     (b) Ne pas proposer φ et retenir un couple (φ,e) tel que $\varphi(e)=0$, ce lui permettra de réfuter φ si, à l'avenir, elle est proposée, et de gagner R.
  2. $\varphi\in CK$, il peut décider de retirer une nouvelle formule ou bien d'entamer un processus de recherche. Si tel est le cas, si $\varphi\in F$, il abandonne immédiatement et tire une nouvelle formule. Si $\varphi\in V$, il essaie de réfuter la théorie en recherchant un résultat d'expérience incompatible jusqu'au temps d'arrêt $\nu(\omega)$. S'il en trouve un, il gagne R.

A chaque pas de temps, l'agent vérifie qu'il n'a pas été accepté dans CK une formule dont il connaît un contre exemple et le cas échéant, exhibe un contre exemple et gagne R.

L'état des connaissances d'une communauté à un moment donné peut être représenté par la figure 5.

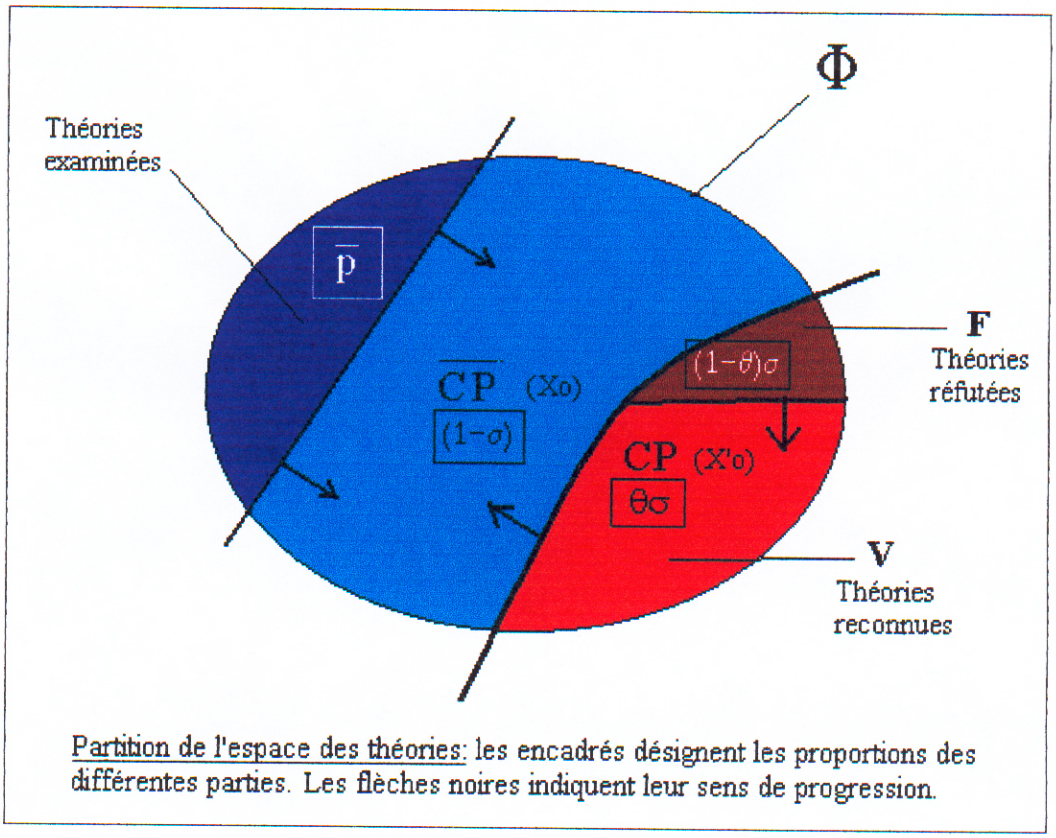

Φ
Théories examinées
$\overline{p}$
$\overline{CP}$ (Xo)
(1−σ)
(1−θ)σ
F
Théories réfutées
CP (X'o)
θσ
V
Théories reconnues
Partition de l'espace des théories: les encadrés désignent les proportions des différentes parties. Les flèches noires indiquent leur sens de progression.

Figure 5:

4.4. Le choix des théories

Le comportement des agents est déterminé par deux types de mécanismes. Localement, quand plusieurs actions sont possibles, un agent choisit l'action associée à l'espérance de gain la plus élevée, selon le principe de l'agent rationnel. L'autre mécanisme majeur, relatif à l'acquisition d'information, concerne les temps d'arrêt caractéristiques des différents processus de recherche. A partir de quel moment l'agent décide-t-il qu'il dispose de suffisamment d'informations à propos d'une théorie pour interrompre le processus de recherche et décider de l'opportunité d'une publication? C'est ce choix de stratégie qui va être primordial pour déterminer l'espérance de gain $G.X$. Notre but sera donc de trouver les temps d'arrêt optimaux vis à vis de la valeur de $G.X$. Le point le plus proche dans l'arbre de décision, concerné par le choix d'un temps d'arrêt est la recherche d'une nouvelle formule. Remarquons en effet que nous avons deux types différents d'information, celle qui donne le statut de la formule (CP ou $\overline{CP}$), celle qui est founie par l'expérience et qui concerne la compatibilité des formules aux tests. Nous avons désigné par X_0 et $\overline{X_0}$ les espérances de gain initiales associées respectivement aux informations CP et $\overline{CP}$, qui correspondent aux gains espérés lorsque l'on commence à tester une nouvelle formule et quand on commence à chercher à réfuter une théorie déjà existante. Le processus est résumé par la figure 6.

La catégorisation des formules en fonction de leur appartenance à CP ou $\overline{CP}$ oblige les agents à adopter des stratégies vis à vis du choix des théories en fonction des valeurs de X_0 et $\overline{X_0}$. Prenons par exemple le cas où $\overline{X_0} < X_0$. Un agent qui tire une formule dans CK est dans le meilleur des cas et donc entame un processus de recherche d'un contre exemple. Par contre, un agent qui tire une formule dans $\overline{CK}$, peut décider de l'abandonner pour tenter d'avoir une formule de CK associée à un gain inital supérieur, au risque de sortir du jeu avec la probabilité α. L'arbre de décision est illustré par la figure 7.

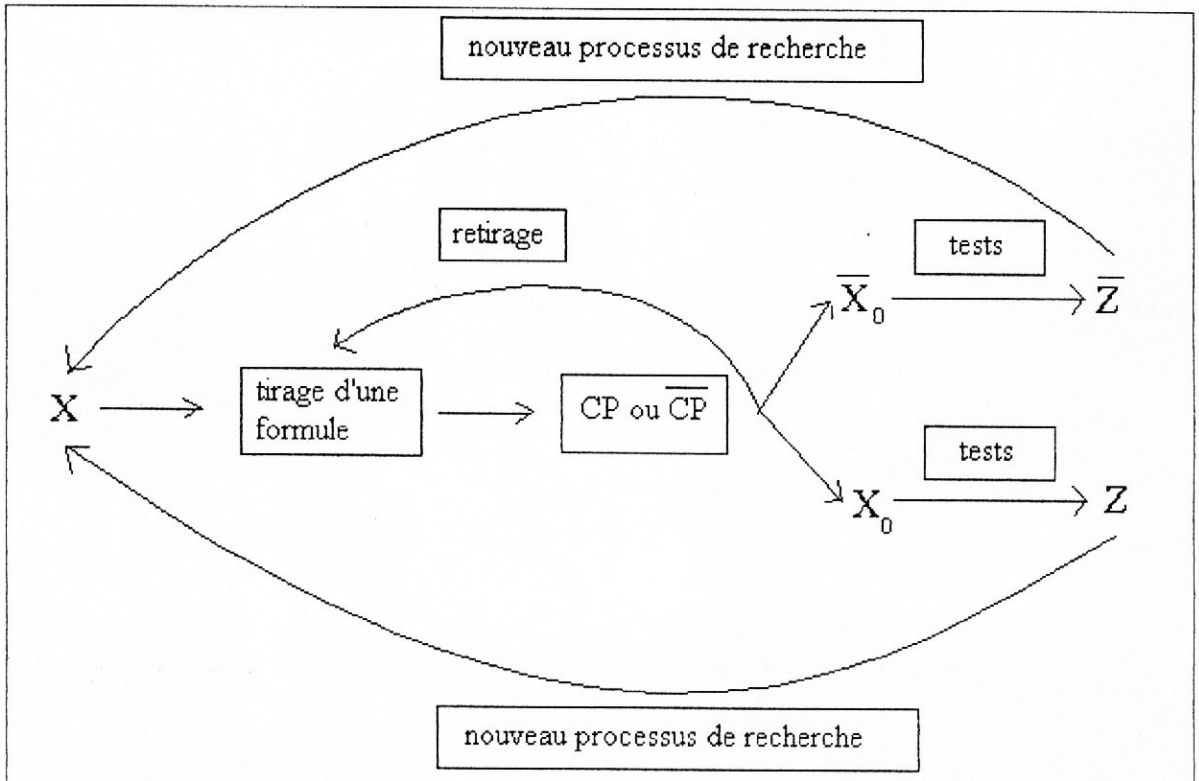
L'agent doit donc adopter des stratégies du type "je consacre n coups à essayer de trouver une formule dans CK. Je commence un processus de recherche au plus tard dans n coups". Il y a donc une phase d'exploration des formules avant d'entamer le processus de recherche même. Cette phase d'exploration est déterminée par le temps d'arrêt ν_n qui optimise X. Ce temps d'arrêt ne dépend des paramètres du modèle qu'à travers X_0 et $\overline{X_0}$ car X est un barycentre de ces deux quantités et peut s'écrire

$$X_{\nu_n,\nu,\overline{\nu}} = d.X_0 + \overline{d}.\overline{X_0}$$

. Les coefficents de ce barycentre dépendent de la manière dont se déroule l'exploration et ont donc des formes mathématiques différentes suivant que l'agent recherche préfère les formules de CP ou les formules de $\overline{CP}$.

si $\overline{X_0} < X_0$:

$$X_{\nu_n,\nu,\overline{\nu}} = \sigma\left(1+\alpha\left(1-\sigma\right)+...+\alpha^n\left(1-\sigma\right)^n\right).X_0+\alpha^n\left(1-\sigma\right)^{n-1}\overline{X_0}$$



Mécanismes de décision: nous pouvons voir trois boucles correspondant à la phase d'exploration des formules et aux phases de confrontation avec les résultats expérimentaux. La taille de ces boucles est déterminée par les temps d'arrêt adoptés pour les différentes phases.

Figure 6:

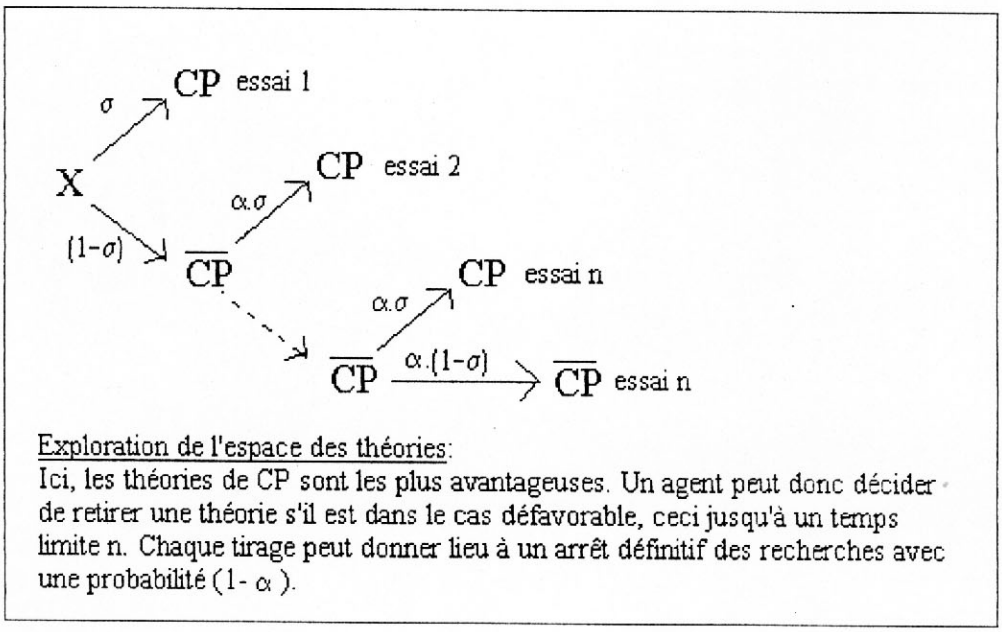


Figure 7:

$$X_{\nu_n,\nu,\overline{\nu}} = \sigma.\frac{1-\alpha^{n+1}(1-\sigma)^{n+1}}{1-\alpha(1-\sigma)}X_0(\nu_n,\nu,\overline{\nu}) + \alpha^n(1-\sigma)^{n-1}\overline{X_0}(\nu_n,\nu,\overline{\nu})$$

Symétriquement si $\overline{X_0} > X_0$:

$$X_{\nu_n,\nu,\overline{\nu}} = \alpha^n\sigma^{n+1}X_0 + (1-\sigma).\frac{1-\alpha^{n+1}\sigma^{n-1}}{1-\alpha\sigma}\overline{X_0}$$

Nous pouvons résumer la situation par le tableau suivant:

	d	$\overline{d}$
$\overline{X_0} < X_0$	$\sigma.\frac{1-\alpha^{n+1}(1-\sigma)^{n+1}}{1-\alpha(1-\sigma)}$	$\alpha^n(1-\sigma)^{n-1}$
$\overline{X_0} < X_0$	$\alpha^n\sigma^{n+1}$	$(1-\sigma).\frac{1-\alpha^{n+1}\sigma^{n+1}}{1-\alpha\sigma}$

(4.1)

Nous pouvons alors définir un paramètre qui sera important pour la suite, la probabilité ℓ pour qu'un agent travaille sur une formule de CK sachant qu'il a entamé un processus de recherche:

$$\ell(\nu_n) = \frac{d}{d+\overline{d}}$$

Il y a donc une phase d'exploration de l'espace des théories qui est déterminée par un temps d'arrêt que nous noterons ν_n. Ce temps d'arrêt indique l'intérêt qu'accorde un agent aux formules de CP et $\overline{CP}$.

Deux autres temps d'arrêt seront importants pour l'étude du comportement de la population. Celui qui guide les processus de recherche dans $\overline{CK}$ et celui qui guide les processus de recherche dans CP, que nous noterons respectivement ν et $\overline{\nu}$.

4.5. Calcul des espérances de gain

Formalisons un peu plus les choix proposés aux agents. Pour un agent caractérisé par son histoire H, un processus de recherche commence par la découverte d'un nouvel élément de H : $(\varphi, \mathcal{E})$. Avant de commencer, l'agent dispose d'une espérance de gain $G.X$ concernant son activité à venir qui dépend des principes qu'il a adoptés pour guider ses recherches. X est la fonction que l'agent va optimiser. Celle-ci dépend des temps d'arrêt choisis et des paramètres de la communauté. Nous supposerons qu'il y a invariance dans le temps, c'est à dire, qu'à tout moment, un agent qui sort d'un processus de recherche se retrouve avec, en ce qui concerne ses gains futurs, l'espérance $G.X$. A l'instant t, l'espérance de gain Z^t associée à un abandon au temps t du processus de recherche en cours se décompose en deux termes, une utilité U^t associée à l'exploitation des informations acquises sur φ,une espérance de gains $G.X$ qui évalue les gains concernant l'activité future de recherche. L'espérance de gain correspondant à une sortie de jeu au temps t pour un tirage φ sera donc de la forme $Z^t = \alpha^t \left(U^t + G.X\right)$. Le terme α^t étant le taux d'actualisation traduisant l'éventualité de la disparition de l'agent avant la fin de ses recherches. Nous voyons alors surgir la principale difficulté du problème posé, à savoir, la stratégie adoptée par un agents dépend du profil des gains associés aux différentes actions, qui lui même dépend de la stratégie adoptée par l'intermédiaire de X. A chaque stratégie différente correspond un jeu différent. L'ensemble des jeux possibles n'est donc pas dénombrable.

Etude pour une recherche dans CP. Un agent qui entame une recherche sur une formule de CP, si elle n'est pas dans F élabore une série d'expériences destinées à tester la théorie. A la première expérience qui contredit les prédictionsde la théorie, il sort du processus de recherche, gagne R et se retrouve avec un avenir caractérisé par l'espérance de gain $G.X$. Tant qu'il ne peut pas réfuter la théorie, l'espérance de gain associée à une sortie de jeu est uniquement $G.X$. L'utilité d'une formule à l'instant t est donc donnée par $U^t = R$ si l'agent a rencontré un contre exemple, et $U^t = 0$ dans le cas contraire. En notant $S_t = \sum_{i=1}^{t} \varphi(\pi(i))$ la variable aléatoire représentant le nombre de fois où la théorie φ a réussi les tests, nous pouvons écrire $U^t = 1_{\{S_t<t\}}.R$. L'espérance de gain associée à l'interruption du processus de recherche dans CP est donc donnée par $Z_\varphi^t = \alpha^t.\left(1_{\{S_t<t\}}.R + G.X\right)$.

Etude pour une recherche dans $\overline{CP}$. Au cours d'une recherche dans $\overline{CP}$ sur une formule φ, le but d'un agent est d'estimer le type φ afin de savoir s'il est avantageux de la proposer. Un agent qui travaille sur φ a , à chaque instant, des croyances sur le type φ qui sont fonction des résultats des expériences qu'ils ont réalisées. Celles-ci sont représentées par une distribution de probabilité sur les types possibles $T_1, ..., T_q, T_\infty$. Nous noterons $p_\varphi(t) = \left(\overline{\mu}_1^t, ..., \overline{\mu}_q^t, \overline{\mu}_\infty^t\right)$ la distribution après t essais sur φ. A chaque fois qu'un agent teste une formule φ il révise ses croyances sur le type de φ en faisant une estimation bayésienne. En notant comme précédemment $S_{t,\varphi,\pi}$, soit en abrégé S_t, la v.a. de Bernouilli

$S_t = \sum_{i=1}^{t} \varphi(\pi(i))$, nous avons alors au temps t:

$$\forall k \in (1, ..., s, \infty), \overline{\mu}_k^t = \frac{B_{t,\theta_k}(S_t).\overline{\mu}_k}{B_{t,1}(S_t).\overline{\mu}_\infty + \sum_{i=1}^{q} B_{t,\theta_i}(S_t).\overline{\mu}_i}$$

Avec: $B_{n,p}(S_t) = C_n^{S_t} p^{S_t}(1-p)^{n-S_t}$

Nous faisons l'hypothèse que les agents connaissent $\left(\overline{\mu}_1, ..., \overline{\mu}_q, \overline{\mu}_\infty\right)$. Ils ont une bonne intuition des différentes catégories de théories qui existent et de leurs fréquences.

A partir de ces croyances sur les types, un agent fait une estimation de l'espérance de gain associée à φ relativement à chaque action possible. A chaque pas de temps, un agent a le choix entre publier, interrompre ses recherches sans publier, et poursuivre ses recherches. Nous avons vu que l'espérance de gain correspondant à une sortie de jeu est de la forme $Z^t = \alpha^t.(U^t + G.X)$. Il nous faut donc déterminer U^t.

U_φ^t est le max entre deux termes:

- gain attendu si l'agent ne propose pas sa formule et garde un contre exemple (dans la mesure où celui-ci en possède un). Ceci dépend de la probabilité pour que, pendant son temps de vie, un autre agent se mette à chercher φ et la propose.
- gain espéré si l'agent propose φ, qui est égal au gain immédiat G diminué de ce qui perdrait s'il était réfuté pondéré par la probabilité pour que cet évènement se produise: $G- R.E^t$(être falsifié de son vivant).

Utilité d'une formule. Associons à chaque type T_i une utilité u_i correspondant au gain espéré si la formule tirée est du type i. Nous avons alors une utilité globale pour une formule après t essais, qui donne le gain espéré si l'agent interrompt le processus de recherche en cours: $U_\varphi^t = u_\infty \mu_\infty^t + \sum_{i=1}^{s} u_i \mu_i^t$. Nous avons alors $Z_\varphi^t = \alpha^t.(u_\infty \mu_\infty^t + \sum_{i=1}^{s} u_i \mu_i^t + G.X)$. Il nous faut donc déterminer les utilités u_i en fonction de $p_\varphi(t)$. Celles-ci seront fonction des risques courus vis à vis d'une éventuelle vérification. Ces risques étant évalués à partir d'anticipations sur les comportements des autres agents nous avons besoin pour les déterminer de définir des variables globales intégrant les comportement des agents dans leur ensemble. Définissons alors le jeu de variables $\left(\overline{f}_k\right)_{k\in(1...,s,\infty)}$, où $\overline{f}_k$ est la probabilité pour qu'un agent qui a travaillé sur une théorie de type k de $\overline{CP}$, décide, en fin de processus de recherche, de ne pas la proposer et de garder un contre exemple. Ces variables représentent la valeur qu'attribue la communauté à une théorie de type k. De la même manière, définissons le jeu de variables $(f_k)_k$, probabilités pour qu'un agent qui cherche un contre exemple à une formule de type k dans CP en trouve effectivement un. Les f_k et les $\overline{f}_k$ sont des variables globales caractérisant les processus de recherche dans CP. et $\overline{CP}$. Ce sont des variables caractéristiques de l'état de la communauté qui intègrent des comportements collectifs. Nous considérerons que les f_k et les $\overline{f}_k$ sont connus de tous. Dans un modèle plus réaliste, ils pourraient être déterminés par l'agent au cours d'un apprentissage.

Remarquons que nous avons une dualité entre les (θ_k) et les $\left(\overline{f}_k\right)$, les premiers relatant la valeur objective des théories, les derniers leur valeur subjective. Pour une interprétation un peu plus parlante,nous pouvons remarquer que $\left(1-\overline{f}_k\right)$ est la probabilité de publication pour une formule de type k lorsque celle-ci est examinée. Nous pouvons maintenant évaluer les risques puis les utilités.

Dans le cas d'une publication:. Nous noterons $\langle\nu\rangle$ le temps moyen de recherche global, intégrant les différents types de processus.

$$\langle\nu\rangle = (1-\ell)\,\langle\overline{\nu}\rangle + \ell.\,\langle\nu\rangle + \langle\nu_n\rangle \tag{4.2}$$

La probabillité pour qu'un agent possède un contre exemple à une formule de type k s'établit par:

- nombre moyen de formules examinées par un agent pendant son temps de vie: $\frac{\langle T\rangle}{\langle\nu\rangle}$
- Probabilité pour qu'il tiré φ dans $\overline{CP}$:

$$\frac{(1-\ell)\,\frac{\langle T\rangle}{\langle\nu\rangle}}{(1-\sigma)\,|\Phi|}$$

- Probabilité pour qu'il ait tiré φ dans $\overline{CP}$ et qu'il ait gardé un contre exemple:

$$\frac{\min(1,\frac{(1-\ell)\frac{\langle T\rangle}{\langle\nu\rangle}}{(1-\sigma)|\Phi|}.\overline{f}_k)}{\left(1-\frac{\frac{\langle T\rangle}{\langle\nu\rangle}}{|\Phi|}\right)+\left(1-\overline{f}_k\right)\frac{\frac{\langle T\rangle}{\langle\nu\rangle}}{|\Phi|}}$$

(formule de Bayes appliquée aux événements "garde un contre exemple" et "ne tombe pas sur φ ")

- Probabilité pour qu'un agent possède un contre exemple:

$$\frac{\min\left(1,\frac{r.(1-\ell)}{(1-\sigma)}.\frac{\langle T\rangle}{\langle\nu\rangle}.\overline{f}_k\right)}{\left(1-\frac{\frac{\langle T\rangle}{\langle\nu\rangle}}{|\Phi|}\right)+\left(1-\overline{f}_k\right)\frac{\frac{\langle T\rangle}{\langle\nu\rangle}}{|\Phi|}}$$

Donc $u_k = G - R.\frac{\min\left(1,\frac{r.(1-\ell)}{(1-\sigma)}.\frac{\langle T\rangle}{\langle\nu\rangle}.\overline{f}_k\right)}{\left(1-\frac{\frac{\langle T\rangle}{\langle\nu\rangle}}{|\Phi|}\right)+\left(1-\overline{f}_k\right)\frac{\frac{\langle T\rangle}{\langle\nu\rangle}}{|\Phi|}}$ si l'agent décide de proposer la formule. Par la suite, nous considérerons que

$$\frac{\frac{\langle T\rangle}{\langle\nu\rangle}}{|\Phi|} \ll 1 \qquad \text{(contrainte 1)}$$

ce qui revient à considérer qu'un agent explore qu'une petite partie des théories possible au cours de sa vie. Nous obtenons alors $u_k = G - R.\min\left(1,\frac{r.(1-\ell)}{(1-\sigma)}.\frac{\langle T\rangle}{\langle\nu\rangle}.\overline{f}_k\right)$

Un agent doit en outre prendre en compte la possibilité de se faire contrôler par un agent effectuant des recherches dans CP . La probabilité pour qu'un agent vérifie à l'avenir une formule φ de type k proposée et trouve un contre exemple s'établit par:

- Probabilité pour qu'un agent donné tire φ dans CP :

$$min\left(1, \frac{\ell.\frac{\langle T\rangle}{\langle\nu\rangle}}{\sigma.\,|\Phi|}\right)$$

- Probabilité pour qu'un agent donné tire φ et trouve un contre exemple:

$$min\left(1, \frac{\ell.\frac{\langle T\rangle}{\langle\nu\rangle}}{\sigma.\,|\Phi|}.f_k\right)$$

- Probabilité pour qu'un agent à l'avenir tire φ et trouve un contre exemple:

$$min\left(1, \frac{\ell.r}{\sigma}\frac{\langle T\rangle}{\langle\nu\rangle}.f_k\right)$$

L'utilité d'une formule de type k associée à une publication est donc:
$u_k = G - R.\left(\min\left(1, \frac{r.(1-\ell)}{(1-\sigma)}.\frac{\langle T\rangle}{\langle\nu\rangle}.\overline{f}_k\right) + min\left(1, \frac{\ell.r}{\sigma}\frac{\langle T\rangle}{\langle\nu\rangle}.f_k\right)\right)$.

Dans le cas où l'agent décide de ne pas publier. Il faut évaluer la probabilité pour que du vivant de l'agent, un autre agent décide de publier la même formule.

- Probabilité pour que φ fasse l'objet d'un tirage dans $\overline{CP}$ à l'avenir :

$$\min(1, \frac{(1-\ell)\,\frac{\langle T\rangle}{\langle\nu\rangle}}{(1-\sigma)\,|\Phi|})$$

- Probabilité pour qu'un agent quelconque propose φ du vivant de l'agent:

$$\min\left(1, \frac{r.\,(1-\ell)}{(1-\sigma)}.\frac{\langle T\rangle}{\langle\nu\rangle}.(1-\overline{f}_k)\right)$$

L'utilité associée à un contre exemple à une formule de type k est alors

$$u_k = R\left(\min\left(1, \frac{r.\,(1-\ell)}{(1-\sigma)}.\frac{\langle T\rangle}{\langle\nu\rangle}.(1-\overline{f}_k)\right)\right)$$

Donc nous avons

$$u_k = G - R.\left(\min\left(1, \frac{r.\ell}{\sigma}\frac{\langle T\rangle}{\langle\nu\rangle}.f_k\right) + \min\left(1, \frac{r.\,(1-\ell)}{(1-\sigma)}.\frac{\langle T\rangle}{\langle\nu\rangle}.\overline{f}_k\right)\right)$$

si l'agent propose la formule et

$$u_k = R.\min\left(1, \frac{r.\,(1-\ell)}{(1-\sigma)}.\frac{\langle T\rangle}{\langle\nu\rangle}.(1-\overline{f}_k)\right)$$

s'il décide de garder un contre exemple.

Pour rendre les formules plus lisibles, nous utiliserons les abréviations:

$$ri(t) = \sum_k \overline{\mu}_k^t . \left(\min\left(1, \frac{r.(1-\ell)}{(1-\sigma)} . \frac{\langle T \rangle}{\langle \nu \rangle} . \overline{f}_k \right) + \min\left(1, \frac{r.\ell}{\sigma} \frac{\langle T \rangle}{\langle \nu \rangle} . f_k \right) \right) \tag{4.3}$$

c'est une estimation du risque associé à la publication d'une formule après t tests.

$$gf(t) = \sum_k \overline{\mu}_k^t . \min\left(1, \frac{r.(1-\ell)}{(1-\sigma)} . \frac{\langle T \rangle}{\langle \nu \rangle} . (1 - \overline{f}_k)\right)$$

c'est une estimation des gains attendus si l'agent décide de garder un contre exemple.

Nous avons alors:

$$Z_\varphi^t = \alpha^t . \left(\max\left[G - R.ri(t), 1_{\{S_t < t\}} . R.gf(t)\right] + \alpha.G.X \right)$$

Espérances de gain initiales. Une fois que l'on dispose de l'expression des gains associés à chaque action, nous pouvons exprimer les espérances de gain initiales relatives à chaque processus comme la moyenne des espérances de gain sur l'ensemble des évènements possibles. Nous avons alors $X_0 = E\left(Z^{\nu(\omega)}\right)$ et $\overline{X}_0 = E\left(\overline{Z}^{\overline{\nu}(\omega)}\right)$

4.6. Calcul des variables globales

Les jeux de variables (f_k) et $\left(\overline{f}_k\right)$ sont calculables à partir des temps d'arrêt ν et $\overline{\nu}$.

- f_k est la probabilités pour qu'un agent qui cherche un contre exemple à une formule de type k dans CK en trouve effectivement un. Cette probabilité s'exprime comme la somme des probabilités d'obtenir un résultat négatif à un test avant le temps d'arrêt $\nu(\omega)$, ω parcourant l'espace des évènements. Nous avons donc $f_k = E\left(1_{\{S_{\nu(\omega)} < \nu(\omega)\}}\right)$
- $\overline{f}_k$ est la probabilité pour qu'un agent qui étudie une théorie de $\overline{CP}$ juge plus favorable de s'arrêter en gardant un contre exemple. Les espérances de gain relatives à une sortie du processus de recherche étaient du type $Z_\varphi^t = \alpha^t.G\left(\max\left[g_1(t), g_2(t)\right] + X_{\nu_n, \nu, \nu_{\overline{\lambda}}}\right)$ ou g_1 correspond à une publication et g_2 au cas où l'agent garde un contre exemple. Les $\overline{f}_k$ sont alors définies par $\overline{f}_k = E^{\Pi}_{\varphi \in T_k}(1_{\{g_1(\overline{\nu}(\Phi)) < g_2(\overline{\nu}(\Phi))\}}(\omega))$.

  Nous voyons alors que le signe de $g_1(\overline{\nu}(\omega)) - g_2(\overline{\nu}(\omega))$ détermine les $\overline{f}_k$.

Nous considèrerons à partir de maintenant que :

$$\overline{p} = \frac{r.(1-\ell)}{(1-\sigma)} . \frac{\langle T \rangle}{\langle \nu \rangle} < 1 \qquad \text{(contrainte 2)}$$

ce qui sera vérifié par les systèmes pour lesquels, étant donné une théorie qui n'a pas encore fait l'objet d'une publication, il y a une probabilité non nulle pour que celle-ci n'ait jamais été examinée par un agent en activité (cf. figure 5). Cette condition peu

contraignante va nous permettre de calculer explicitement les $\overline{f}_k$. En effet, nous obtenons ainsi:

$$\begin{aligned} g_1(\nu(\omega)) - g_2(\nu(\omega)) &= G - R.\sum_k \overline{\mu_k}^t.\left(min\left(1, \frac{r.\ell}{\sigma}\frac{\langle T\rangle}{\langle\nu\rangle}.f_k\right) + \overline{p}.\overline{f}_k\right) \\ &\quad -1_{\{S_t<t\}}R.\left(\overline{p} - \sum_k \overline{\mu}_k^t.\overline{p}.\overline{f}_k\right). \end{aligned}$$

Quand $S_{\nu(\omega)} < \nu(\omega)$, le choix de l'agent dépend donc du signe de

$$\varepsilon = G - R.\overline{p} - R.\sum_k \overline{\mu_k}^t.min\left(1, \frac{r.\ell}{\sigma}\frac{\langle T\rangle}{\langle\nu\rangle}.f_k\right)$$

. Ceci dépend des $\overline{f}_k$ et des f_k. Nous pouvons cependant délimiter des zones dans l'espace des paramètres. En effet, nous pouvons faire l'encadrement $G - 2.R \leq \varepsilon \leq G - R.\overline{p}$. Ceci délimite trois zones:

1. $G \geq 2.R$, $g_1(\nu(\omega)) - g_2(\nu(\omega)) \geq 0$ quelque soit l'évènement considéré. Les agents décident alors de publier quelque soit les résultats des tests. Les f_k sont alors tous nuls. Les agents proposent n'importe quel type de théorie. Cette zone de l'espace des paramètres correspond donc à un comportement non scientifique de la part des agents.
2. $G < R.\overline{p}$, un agent ne propose jamais une théorie qui a échouée à l'un des tests. Cette zone de l'espace des paramètres correspond à une communauté où les agents sont honnêtes.
3. $R.\overline{p} \leq G \leq 2.R$, suivant leur histoire les agents peuvent proposer une théorie qui s'est révélée fausse s'ils jugent qu'elle n'est pas trop fausse.

Nous pouvons résumer cette partition de l'espace des paramètres par la figure 8.

Nous allons maintenant nous intéresser au cas 2) exposé précédemment correspondant à une communauté d'agents honnêtes. Ce caractère honnête peut d'ailleurs être imposé par la communauté, qui discrédite un agent qui propose une formule fausse en connaissance de cause, ce qui mettrait en jeu la réputation d'un individu.

5. Comportements scientifiques

Nous allons maintenant étudier le cas 2) où $G < R.\overline{p}$. Nous imposerons cette contrainte dans toute la suite du mémoire.

$$G < R.\overline{p} \qquad \text{(contrainte 3)}$$

Sous cette hypothèse, nous allons voir que nous pouvons analyser beaucoup plus en détail certains aspects de ces systèmes. A cette fin, nous allons pousser plus en avant

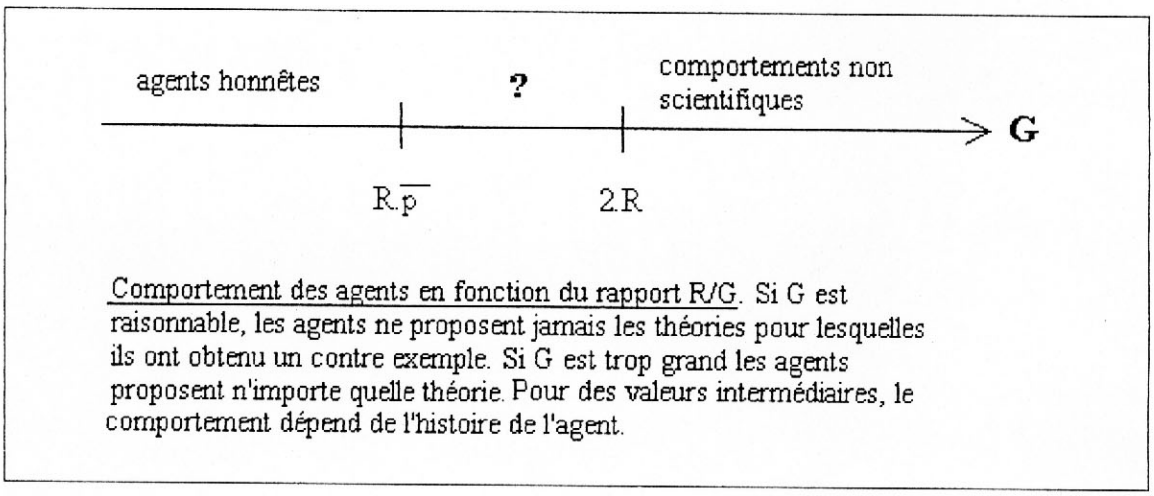


Figure 8:

la formalisation en définissant précisément l'espace des évènements. Puis nous verrons que nous pouvons caractériser les temps d'arrêt optimaux et du même coup les variables globales (f_k) et $\left(\overline{f}_k\right)$.

5.1. L'espace des évènements

L'espace des évènement est discret et est somme directe de deux sous espaces Φ et Π munit des tribus discrètes et des probabilités uniformes. Nous avons d'autre part une fonction

$$\begin{array}{rcl} S & : & \mathbb{N} \times (\Phi \otimes \Pi) \to \mathbb{N}_t \\ t, \varphi, \zeta & \to & S_t(\varphi, \zeta) = \sum_{i \leq t} \varphi(\zeta(i)) \end{array}$$

qui donne le nombre de fois où une théorie donnée φ a passé les tests avec succès avant le temps t. La statistique $S(\varphi, \zeta, t)$ étant exhaustive pour ce modèle, elle contient toute l'information supplémentaire obtenue au temps t. En notant $\overline{h} : \Phi_s \to \{1..., q\}$ la fonction qui à chaque théorie associe son type, définissons la relation d'équivalence $\approx$ sur $\Phi \otimes \Pi$ par $(\varphi, \zeta) \approx (\varphi', \zeta') \Leftrightarrow (h(\varphi') = h(\varphi)) \wedge (\forall i, \varphi'(\zeta'(i)) = \varphi(\zeta(i)))$ et posons $\tilde{\Omega} \otimes \tilde{\Pi} = (\Phi \otimes \Pi)_{/\approx}$. Nous pouvons alors définir:

$\forall\ (t, s) \in^2, s \leq t, \tilde{S}_t^{-1}(s) = \left\{\left(\tilde{\Phi}, \tilde{\zeta}\right) \mid \tilde{\Phi}\left(\tilde{\zeta}\right) = s\right\}$

Au cours d'un processus de recherche, les agents catégorisent les théories en fonction des résultats des expériences, c'est à dire en fonction des valeurs de S_t. Pour une recherche dans V (resp. $\overline{CP}$) la famille de tribus attachée à notre problème est alors la famille indexée par t :

$\mathcal{B}^t = \sigma\left(\cup_{m \leq t}\left\{V \cap \tilde{S}_m^{-1}(s) \mid s \leq m\right\}\right)$ ($\sigma(A)$ désigne la tribu engendrées par A)

(resp. $\overline{\mathcal{B}}^t = \sigma\left(\cup_{m\leq t}\left\{\overline{CK}\cap\tilde{S}_m^{-1}(s)\mid s\leq m\right\}\right)$

Pour se convaincre que ce sont bien les tribus associées au problème posé, nous pouvons donner une représentation graphique pour par exemple, le cas d'une recherche dans $\overline{CP}$. Graphiquement, au temps t,pour chaque type représenté par une formule cannonique φ_k, un agent sélectionne dans $\overline{CP}\otimes\Pi$ l'ensemble des couples (φ_k,ζ) tels que $\sum_{i\leq t}\varphi(\zeta(i)) = S_t$. La mesure de cet ensemble représente la probabilité $\mu_k.B_{t,\theta_k}(S_t)$ qu'une formule de $\overline{CP}$ soit de type k et fournisse la somme S_t après t essais. La probabilité que l'agent attribue à un type est alors $\mu_k^t = \frac{P(\{\varphi_k,\zeta\})}{P(\overline{CP})}$ ce qui nous redonne l'expression des μ_k^t obtenue à l'aide de la formule de Bayes.

5.2. Les temps d'arrêt optimaux

Temps d'arrêt optimal pour une recherche dans $\overline{CP}$. L'espace des évènements considéré est ici $\overline{CP}\otimes\Pi$. Pour un processus de recherche dans $\overline{CP}$ sur une théorie φ nous avons au temps t l'espérance de gain

$$Z_\varphi^t = \alpha^t.\left(\max\left[G - R.ri(t), 1_{\{S_t<t\}}.R.gf(k)\right] + GX\right)$$

Remarquons que Z_φ^t dépend de X et donc des temps d'arrêt $\nu_n, \overline{\nu}$ et ν. Nous le noterons $Z_{\nu_n,\overline{\nu},\nu}^t$ ou Z^t quand il y aura ambiguité. Nous allons essayer de trouver la forme que doivent avoir le temps d'arrêt $\overline{\nu}$ pour que le triplet $(\nu_n,\nu,\overline{\nu})$ puisse être optimal, indépendamment de la forme de ν_n et ν. Nous supposerons donc que les temps d'arrêt ν et ν_n sont fixés de manière quelconque. Nous allons nous intéresser à la borne supérieure de $X(\nu_n,\nu,\overline{\nu})$ quand $\overline{\nu}$ varie. L'espace des temps d'arrêt étant non dénombrable, nous allons rechercher une classe dénombrable de temps d'arrêt dominants sur laquelle il sera plus aisé de faire une optimisation. Pour comprendre la structure des temps d'arrêts nous allons nous aider du lemme:

LEMMA 1. *Nous rappelons que la variable aléatoire $S(t)$ donne le nombre de tests corroborant au temps t. Soit $\overline{\nu}$ un temps d'arrêt, s'il existe $\omega\in\overline{CP}\otimes\Pi$ tel que $\overline{\nu}(\omega) = t$ et $S_t(\omega) = t$ alors pour tout ω' tel que $S_t(\omega') = t$, $\overline{\nu}(\omega') = t$. Autrement dit, il n'y a qu'une seule façon de s'arrêter sur une suite de 1.*

Proof. . La statistique (m, S_m) étant exhaustive pour le modèle considéré, le temps d'arrêt $\overline{\nu}$ ne dépend de ω qu'à travers la suite $((m, S_m(\omega)))_{m>0}$. Soit donc $\omega'\in\overline{CP}\otimes\Pi$ tel que $S_t(\omega') = t$.Comme $\overline{\nu}(\omega) = t$, et que $S_t(\omega) = t$, $((m, S_m(\omega')))_{m\leq t} = ((m, S_m(\omega)))_{m\leq t}$. ω et ω' sont donc dans la même classe d'équivalence à l'instant t. Ils sont donc indistinguables par $\mathcal{B}^t$ et d'après les propriétés des temps d'arrêt, nous avons donc $\omega'\in\{\overline{\nu} = t\}$. ■

Considérons alors le temps d'arrêt défini par:

$$\nu_\infty(\omega) = \inf_{t>0}\left\{t\mid\tilde{S}_t(\tilde{\omega}) < t\right\}$$

Ceci est bien un temps d'arrêt puisque $\forall n, \{\nu_\infty = n\} = \overline{CP} \cap \tilde{S}_n^{-1}(n-1) \cap \tilde{S}_{n-1}^{-1}(n-1) \in \overline{\mathcal{B}}^n$. Ce temps d'arrêt est infini sur la portion d'espace $\overline{CP} \cap T_\infty \otimes \Pi$ et est fini presque sûrement sur son complémentaire. Par la suite nous noterons B_t le sous ensemble de l'espace des évènements $\overline{CP} \cap \tilde{S}_t^{-1}(t-1) \cap \tilde{S}_{t-1}^{-1}(t-1)$. Cette famille d'ensembles est associée naturellement à la partition $\left(\overline{CP} \cap T_\infty \otimes \Pi\right) \otimes_{t \geq 0} B_t$ de $\overline{CP} \otimes \Pi$, catégorisation de l'espace des évènements en fonction de la date d'apparition du premier résultat nul. Nous pouvons alors définir la famille de temps d'arrêt finis $(\nu_\lambda)_{\lambda \in \mathbb{N}} = \Lambda$ telle que

$$\nu_\lambda = \nu_\infty \wedge \lambda$$

où λ est le temps d'arrêt constant. Cette famille est clairement une famille de temps d'arrêts. Nous allons montrer qu'elle forme une famille de temps d'arrêt dominants pour un processus de recherche dans $\overline{CP}$: pour tout temps d'arrêt $\overline{\nu}$, il existe $\lambda \in \mathbb{N}$ tel que $X(\nu_n, \nu, \overline{\nu}) \leq X(\nu_n, \nu, \nu_\lambda)$.

PROPOSITION 2. *Le temps d'arrêt optimal pour une recherche dans $\overline{CP}$ appartient à l'ensemble Λ*

Proof. . Cette démonstration étant assez longue, elle figure en annexe afin de clarifier le mémoire. ■

Cette proposition a une conséquence très importante:

L'espérance de gain optimale si elle existe, est de la forme $X_{\nu_n, \nu, \nu_{\overline{\lambda}}}$

Temps d'arrêt optimal pour une recherche dans CP. Lors d'une recherche dans CP, les agents recherchent des contre-exemples ou des expériences incompatibles avec la théorie. Si la théorie tirée s'avère être dans F, il n'y a pas de recherche possible et la recherche s'arrête immédiatement. Si par contre, elle appartient à V,étant donné que l'espérance de gain n'augmente plus à partir du moment où l'on possède un contre exemple, les stratégies les meilleures sont du type "je m'arrête dès que j'ai trouvé un contre exemple. Je sors au plus tard au bout de λ pas de temps". La stratégie optimale sur V est donc encore dans Λ.

PROPOSITION 3. *L'espérance de gain optimale est de la forme $X_{\nu_n, \nu_\lambda, \nu_{\overline{\lambda}}}$*

Proof. . Nous obtenons cette proposition en regroupant les résultats précédants ■

.

Nous pouvons remarquer que les stratégies d'exploration de l'espace des formules sont similaires aux stratégies de Λ, ce qui justifie la notation.

5.3. Expression des variables globales

Maintenant que nous connaissons l'allure des stratégies optimales, nous pouvons établir l'expression des variable f_k et $\overline{f}_k$ qui leur sont associées. Nous considérerons donc maintenant des systèmes déterminé par des stratégies du type $(\nu_n, \nu_\lambda, \nu_{\overline{\lambda}})$.

Calcul des f_k. f_k est la probabilités pour qu'un agent qui cherche un contre exemple à une formule de type k dans CP en trouve effectivement un. Comme le temps d'arrêt est dans Λ, cette probabilité s'exprime en fonction de λ comme la somme des probabilités d'obtenir un résultat nul au temps t, pour $t \leq \lambda$:

$$f_k = \sum_{t=1}^{\lambda} \alpha^t . (1 - \theta_k) \theta_k^{t-1}$$

que nous pouvons approcher par $f_k = 1 - \left(1 - \frac{1}{2^k}\right)^\lambda$

Calcul des $\overline{f}_k$. Les $\overline{f}_k$ sont définies par $\overline{f}_k = E^{\Pi}_{\varphi \in T_k}(1_{\{g_1(\nu(\omega)) < g_2(\nu(\omega))\}}(\omega))$. Le temps d'arrêt étant de la forme $\nu_{\overline{\lambda}} \in \Lambda$, les cas où un agent décide de publier se réduit aux cas où une théorie a réussi tout les tests jusqu'au temps $\overline{\lambda}$. L'expression des $\overline{f}_k$ en découle:

$$\overline{f}_k = \sum_{t=1}^{\overline{\lambda}} \alpha^t . (1 - \theta_k) \theta_k^{t-1}$$

5.4. Expression des espérances de gain initiales

Maintenant que nous connaissons l'allure des stratégies optimales, nous pouvons calculer les espérances de gain initiales relatives à chaque processus de recherche en fonction de temps d'arrêt fixés $(\nu_n, \nu_\lambda, \nu_{\overline{\lambda}})$. Nous ferons plus tard varier $(\nu_n, \nu_\lambda, \nu_{\overline{\lambda}})$ pour trouver les stratégies optimales.

Pour une recherche dans CP, nous avons :

$$X_0 = E\left(Z^{\nu_\lambda(\omega)}\right) = E\left(1_V . Z^{\nu_\lambda}\right) + E\left(1_F . Z^{\nu_\lambda}\right)$$

$$\begin{aligned} X_0 &= \theta . \textstyle\sum_k \mu_k \left(\sum_{t=1}^{\lambda} \alpha^t . (1 - \theta_k) . \theta_k^{t-1} \left(R + \alpha . G . X_{\nu_n, \nu_\lambda, \nu_{\overline{\lambda}}}\right) + \alpha^{\lambda - 1} . \theta_k^{\lambda} G . X_{\nu_n, \nu_\lambda, \nu_{\overline{\lambda}}}\right) \\ &\quad + (1 - \theta) . G . \alpha . X_{\nu_n, \nu_\lambda, \nu_{\overline{\lambda}}} \end{aligned}$$

C'est la somme sur tous les évènements possibles, des gains obtenu grâce à la découverte d'un contre exemple avant le temps λ lorsque l'on se donne le temps d'arrêt ν_λ.

De même, pour une recherche dans $\overline{CP}$, l'espérance associée au triplet $(\nu_n, \nu_\lambda, \nu_{\overline{\lambda}})$ est donnée par $\overline{X}_0 = E\left(Z^{\nu_{\overline{\lambda}}(\omega)}\right)$:

$$\begin{aligned} \overline{X}_0 &= \overline{\mu_\infty} . \alpha^{\overline{\lambda}} . \left(G - R.ri\left(\overline{\lambda}\right) + \alpha . G . X_{\nu_n, \nu_\lambda, \nu_{\overline{\lambda}}}\right) \\ &\quad + \textstyle\sum_{k=1}^{q} \overline{\mu_k} . \left(\begin{array}{c} \sum_{t=1}^{\overline{\lambda}} \alpha^t (1 - \theta_k) . \theta_k^{t-1} \left(R.gf(t) + \alpha . G . X_{\nu_n, \nu_\lambda, \nu_{\overline{\lambda}}}\right) \\ + \alpha^{\overline{\lambda}} \left(\theta_k^{\overline{\lambda}} . \left(\left(G - R.ri\left(\overline{\lambda}\right)\right) + G . \alpha . X_{\nu_n, \nu_\lambda, \nu_{\overline{\lambda}}}\right)\right) \end{array} \right) \end{aligned}$$

C'est la somme sur $\Phi \otimes \Pi$ des espérances de gains associées au temps d'arrêt $\nu_{\overline{\lambda}}$. Pour clarifier notons:

- $\overline{\gamma}_k = \sum_{t=1}^{\overline{\lambda}} \alpha^t \left(1-\theta_k\right).\theta_k^{t-1}$, nous avons alors $\overline{\gamma}_k = \alpha.\left(1-\theta_k\right).\frac{1-(\alpha.\theta_k)^{\overline{\lambda}-1}}{1-\alpha.\theta_k}$
- $\overline{\delta}_k = \alpha^{\overline{\lambda}}.\theta_k^{\overline{\lambda}}$

Nous avons alors:

$$\begin{aligned}\overline{X}_0 &= G.X_{\nu_n,\nu_\lambda,\nu_{\overline{\lambda}}}.\left(\mu_\infty.\alpha^{\overline{\lambda}+1} + \left(\alpha.\sum_{k=1}^s \mu_k.\left(\overline{\gamma}_k + \overline{\delta}_k\right)\right)\right) \\ &\quad +G.\left(\mu_\infty.\alpha^{\overline{\lambda}} + \sum_{k=1}^s \mu_k.\overline{\delta}_k\right) \\ &\quad +R.\left(\begin{array}{c}\sum_{k=1}^s \overline{\mu}_k.\left(\begin{array}{c}\sum_{t=1}^{\overline{\lambda}} \alpha^t \frac{1}{2^k}\left(1-\frac{1}{2^k}\right)^{t-1} gf(t) \\ -\overline{\delta}_k.\sum_k \overline{\mu}_k^t.ri\left(\overline{\lambda}\right)\end{array}\right) \\ -\overline{\mu_\infty}.\alpha^{\overline{\lambda}}.ri\left(\overline{\lambda}\right)\end{array}\right)\end{aligned}$$

Nous pouvons alors remarquer que de par la forme des gains Z, G se met naturellement en facteur dans les expressions des gains. Quitte a remplacer $\overline{X}_0$ par $G.\overline{X}_0$ et X_0 par $G.X_0$, G est en facteur partout et nous pouvons exprimer le problème en remplaçant le couple (G,R) par la rapport $\frac{R}{G}$.

Notons $X\left(n,\lambda,\overline{\lambda}\right)$ pour $X_{\nu_n,\nu_\lambda,\nu_{\overline{\lambda}}}$. Les espérances de gain associées aux temps d'arrêt $(\nu_n,\nu_\lambda,\nu_{\overline{\lambda}})$ s'obtiennent en résolvant le système:

$$(1):\left\{\begin{array}{r} X_0 = u.X\left(n,\lambda,\overline{\lambda}\right) + v \\ \overline{X_0} = \overline{u}.X\left(n,\lambda,\overline{\lambda}\right) + \overline{v} \\ X\left(n,\lambda,\overline{\lambda}\right) = \overline{d}(n)\,\overline{X_0} + d(n).X_0 \end{array}\right. \tag{5.1}$$

avec:

$$\begin{array}{|l|}\hline u(\theta,\lambda) = \theta.\sum_k \mu_k\left(\sum_{t=1}^{\lambda} \alpha^{t+1}.\left(1-\theta_k\right).\theta_k^{t-1} + \alpha^{\lambda-1}.\theta_k^{\lambda}\right) + \alpha.(1-\theta) \\ \hline v\left(\frac{R}{G},\theta,\lambda\right) = \frac{R}{G}.\theta.\sum_k \mu_k \sum_{t=1}^{\lambda} \alpha^t.\left(1-\theta_k\right).\theta_k^{t-1} \\ \hline \overline{u}\left(\theta,\overline{\lambda}\right) = \frac{R}{G}.\theta.\sum_k \mu_k \sum_{t=1}^{\lambda} \alpha^t.\left(1-\theta_k\right).\theta_k^{t-1} \\ \hline \overline{v}\left(\frac{R}{G},r,\sigma,\theta,n,\overline{\lambda}\right) = .\mu_\infty.\alpha^{\overline{\lambda}} + \sum_{k=1}^s \mu_k.\overline{\delta}_k \\ \qquad +R.\left(\begin{array}{c}\sum_{k=1}^s \overline{\mu}_k.\left(\begin{array}{c}\sum_{t=1}^{\overline{\lambda}} \alpha^t \frac{1}{2^k}\left(1-\frac{1}{2^k}\right)^{t-1} gf(t) \\ -\overline{\delta}_k.\sum_k \overline{\mu}_k^t.ri\left(\overline{\lambda}\right)\end{array}\right) \\ -\overline{\mu_\infty}.\alpha^{\overline{\lambda}}.ri\left(\overline{\lambda}\right)\end{array}\right) \\ \hline\end{array} \tag{5.2}$$

Nous en déduisons le système linéaire qui nous permet d'obtenir les valeurs de X_0 et $\overline{X_0}$:

$$\left\{\begin{array}{l}\left(1-\overline{u}.\overline{d}(n)\right)\overline{X}_0 - \overline{u}.d(n).X_0 = \overline{v} \\ -u.\overline{d}(n).\overline{X}_0 + \left(1-u.d(n)\right)X_0 = v\end{array}\right.$$

En posant $\Delta = \left(1 - \overline{u}.\overline{d}(n)\right).(1 - u.d(n)) - u.\overline{d}(n).\overline{u}.d(n)$,

$$\Delta = 1 - \left(\overline{u}.\overline{d}(n) + u.d(n)\right) > 0$$

nous avons:

$X_0 = \frac{\overline{v}.(1-u.d(n))+v.\overline{u}.d(n)}{\Delta}$ et $X_0' = \frac{\overline{v}.u.\overline{d}(n)+v.(1-\overline{u}.d(n))}{\Delta}$ où il nous faut prendre les valeurs de $d(n)$ et $\overline{d}(n)$ compatibles avec les valeurs de X_0 et $\overline{X_0}$. Nous obtenons ensuite $X_{\nu_n,\nu_\lambda,\nu_{\overline{\lambda}}}$ en faisant $X_{\nu_n,\nu_\lambda,\nu_{\overline{\lambda}}} = \overline{d}(n)\overline{X}_0 + d(n).X_0$.

5.5. Stratégie optimale pour un équilibre statique

Nous avons montré que dans le cas où $G < R.\overline{p}$, les triplets $(\nu_n, \nu_\lambda, \nu_{\overline{\lambda}})$ forment une classe de stratégies dominantes pour des équilibres statiques. Les agents cherchant à optimiser leur espérance totale de gain X, ils adoptent une stratégie du type $(\nu_n, \nu_\lambda, \nu_{\overline{\lambda}})$. S'il existe un équilibre pour ce système, à l'équilibre, les f_k et $\overline{f}_k$ sont constants et tous les agents adoptent la stratégie optimale associée à ces jeux de constantes. Comme les stratégies optimales sont du type $(\nu_n, \nu_\lambda, \nu_{\overline{\lambda}})$, les f_k et $\overline{f}_k$ dépendent de λ et de $\overline{\lambda}$ et ont les expressions calculées précédemment. Dès lors, X se calcule en fonction de λ et de $\overline{\lambda}$ grâce aux expressions ci-dessus. L'équilibre statique est donc déterminé, s'il existe, comme étant le maxmum de $X_{\nu_n,\nu_\lambda,\nu_{\overline{\lambda}}}$ quand n, λ et $\overline{\lambda}$ varient sur $\mathbb{N}$. Il est cependant possible que la dynamique du système résultant de comportements individuels optimaux ne soit pas un équilibre statique mais qu'il y ait des fluctuations autour d'une position d'équilibre. Supposons en effet que les agents initialement adoptent tous, la stratégie correspondant à un équilibre statique optimal. Les f_k et $\overline{f}_k$ sont alors fixés par rapport à cette stratégie. Des agents entamant un nouveau processus de recherche pourraient alors profiter de ces valeurs localement constantes pour augmenter ou diminuer leurs temps de recherche ce qui entrainerait, à terme, une réévaluation des f_k et $\overline{f}_k$. Sous réserve que ces fluctuations soient petites, nous pouvons trouver une bonne approximation du comportement de la société en considérant que l'on a affaire à des équilibres statiques. Nous trouvons alors les stratégies optimales associées à des distributions données de (μ_k) et $(\overline{\mu}_k)$ en trouvant le maximum $X_{\nu_n,\nu_\lambda,\nu_{\overline{\lambda}}}$ en n, λ et $\overline{\lambda}$. Le taux d'actualisation α que nous avons introduit nous permet d'effectuer cette recherche sur une partie bornée de $\mathbb{N}^3$ et nous assure de l'existence d'un optimum.