

## Part III

# Application à un système logique

### 6. Introduction

Nous allons maintenant concrétiser l'étude précédente en considérant un espace particulier de théories. Nous allons prendre comme espace des théories un espace de formules logiques  $\Phi$  sur un ensemble de variables propositionnelles  $p_1, \dots, p_n$ . Nous rappelons que cette formalisation ne fait aucune assertion vis à vis de la forme des théories considérées qui peuvent être sous forme linguistique, sous forme de schémas ou d'équations. La seule chose exigée est qu'elles puissent être réfutées par l'expérience. Nous avons choisi une formalisation logique pour une raison simple: les opérateurs  $\wedge$  (*et*) et  $\vee$  (*ou*) permettent de composer les théories de manière naturelle en agissant sur l'ensemble des énoncés de base par union et intersection. Dans une optique de description du monde par sélection des énoncés de base, ces opérateurs sont donc essentiels. L'ensemble des valuations de ces formules sera noté  $\mathcal{E}$ . Un point de  $\mathcal{E}$  sera interprété comme un énoncé de base. L'ensemble des énoncés permis par une théorie  $\varphi$  sera alors l'ensemble des énoncés de base sur lesquels la théorie vaut 1. L'ensemble des falsificateurs virtuels de la théorie sera l'ensemble des énoncés sur lesquels la théorie vaut 0. L'ensemble des contre-exemples à la théorie sera l'ensemble des énoncés correspondant à une description de phénomènes réels, sur lesquels la théorie vaut 0. A la différence des modèles logiques kripkéens, où une certaine fonction d'évaluation de formule (par exemple  $(p_1 = 1, p_2 = 0, \dots, p_n = 1)$ ) est interprétée comme un monde possible, chaque booléen est ici interprété comme un fait potentiellement observable dans le monde. Un monde possible, identifié par l'ensemble des faits qui y sont observables, sera alors un ensemble de valuations. Le but des agents est alors de trouver les formules logiques vraies sur toute valuation du monde réel. Plus précisément, nous savons que dans le cadre du calcul propositionnel, à chaque ensemble de valuations, nous pouvons associer une formule logique dont le domaine de validité est exactement cet ensemble. Si le monde réel est caractérisé par un ensemble  $\mathcal{E}^{réel}$ , à cet ensemble correspond une formule logique  $\Xi$  telle que  $\mathcal{E}_{\Xi} = \mathcal{E}^{réel}$ . Nous pouvons dire que  $\Xi$  décrit exactement le monde réel en ce sens que son extension est composée de tout les faits observables dans le monde et d'eux seuls. Nous pouvons écrire  $\Xi$  sous forme normale conjonctive  $C_1 \wedge C_2 \wedge \dots \wedge C_n$  avec  $C_k$  du type  $q_1 \vee \dots \vee q_r$ ,  $q_i \in (p_1, \dots, p_n, \neg p_1, \dots, \neg p_n)$ . Le but des agents est alors de trouver les clauses que comporte la théorie générale.  $CP$  est alors composé de clauses *disjonctives* qui imposent chacune des contraintes au monde réel en interdisant certaines zones de l'ensemble des possibles. La

représentation  $\chi$  que de fait la communauté du monde est alors la conjonction de toutes les théories de  $\chi = CP \wedge_{\varphi \in CP} \varphi$ . C'est donc une formule sous forme normale conjonctive. La communauté révisé donc ses connaissances en testant des clauses de  $\chi$  et en les modifiant si besoin est.

Nous considérerons donc comme espace  $\Phi$  des théories *l'ensemble des formules logiques sous forme disjonctive* comportant  $s$  éléments sur un ensemble de variables propositionnelles  $p_1, \dots, p_n$ , ces éléments pouvant par ailleurs être répétés. Pour un élément  $q_i \in (p_1, \dots, p_n, \neg p_1, \dots, \neg p_n)$  nous noterons  $\bar{q}_i$  son complémentaire. Un ensemble  $\{q_1, \dots, q_r\} \subset (p_1, \dots, p_n, \neg p_1, \dots, \neg p_n)$  sera dit consistant s'il ne contient pas de paire de littéraux complémentaires. Faisons quelques remarques visant à interpréter sur les constantes  $n$  et  $s$ . Nous avons fait dans les parties précédentes une distinction entre falsificateurs virtuels et contre-exemples. Nous avons vu qu'une théorie est d'autant plus informative que le nombre de ses falsificateurs virtuels est grand et que sa probabilité logique est faible. Elle restreint d'autant plus l'ensemble des possibles qu'elle interdit plus de choses. Dans le cas de l'espace  $\Phi$ , la probabilité logique d'une formule croît proportionnellement avec sa taille. La taille d'une formule disjonctive est donc un indice de son degré d'informativité. Nous ne prendrons pas en compte, dans ces études, ce degré d'informativité. Il serait pourtant essentiel si l'on veut pouvoir trouver les clauses  $C_1 \wedge C_2 \wedge \dots \wedge C_n$  qui correspond à l'ensemble *minimal* susceptible de décrire le monde. Nous verrons plus tard comment  $n$  peut être interprété comme un indice de l'avancement technologique de la communauté. Nous considérerons donc que le agents ne se préoccupent pas de savoir si les formules qu'ils proposent sont assez précises, mais seulement si elle sont vraies dans le monde réel. Les variables pertinentes dans ce problème pour décrire les théories sont alors les  $\theta_k$  et seuls les énoncés de base décrivant le monde réel seront importants. De ce fait, nous considérerons que les énoncés de base décrivant le monde réel, ensemble que nous paramètrons par  $\mathcal{E}^{réel} = [0, 1]^n$ . Une théorie sera alors de la forme  $q_1 \vee \dots \vee q_s$  avec  $q_i \in (p_1, \dots, p_n, \neg p_1, \dots, \neg p_n)$ . Nous désignerons indifféremment les théories par leur formule ou par l'ensemble des variables propositionnelles qui les composent.

Ce qui nous intéressera désormais, c'est la manière dont des théories fausses peuvent être tenues pour vraies pendant un temps en accédant à  $CP$ . Nous nous intéresserons notamment à la manière dont les agents gèrent le compromis exploration/exploitation en fonction de  $G$  et  $R$  et des différents états de l'ensemble des connaissances, et les répercussions que cela a sur la constitution de  $CP$ .

## 7. Caractéristiques du modèle

### 7.1. Les théories

Une théorie est de la forme  $(q_1, \dots, q_s)$  avec  $q_i \in (p_1, \dots, p_n, \neg p_1, \dots, \neg p_n)$ . Il peut y avoir des termes répétés. Si l'ensemble caractérisant la théorie n'est pas consistant, la théorie

en question est une tautologie et  $\theta = 1$ . Nous noterons  $T_\infty$  la classe des tautologies. Pour ce qui concerne les ensembles consistants, nous pouvons les classer par types  $T_1, \dots, T_s$ , en fonction du nombre de variables distinctes qui y figurent. Nous allons montrer qu'il y a une relation entre le type d'une formule et le paramètre  $\theta_\varphi$  de la variable de Bernoulli  $\varphi(v_\omega)$ .

LEMMA 4. Pour une formule  $\varphi$  de type  $T_k$ ,  $k \in (1, \dots, s)$ ,  $\theta_\varphi = 1 - \frac{1}{k}$

**Proof.** . Si  $\varphi = q_1 \vee q_2 \vee \dots \vee q_s$ , est de type  $T_k$ , et si  $(l_1, \dots, l_k)$  est l'ensemble consistant correspondant alors  $v \in \Phi_2$ ,  $v(\varphi) = 0 \iff \forall i \in (1, \dots, k), v(l_i) = 0$  si le tirage de  $v$  sur  $\Phi$  est uniforme, la probabilité pour que  $v(\varphi) = 0$  est donc  $\left(\frac{1}{2}\right)^k$  et  $\theta_\varphi = 1 - \frac{1}{2^k}$ . ■

Il s'agit donc maintenant d'estimer la probabilité de chaque type pour un tirage aléatoire de loi uniforme sur  $\Phi$ .

LEMMA 5. La probabilité  $\mu_k$  du type  $k$  est  $\mu_k = \frac{2^k C_n^k C_{s-1}^{k-1}}{\sum_{i=1}^s C_{2n}^i C_{s-1}^i}$ .

**Proof.** . Une formule de type  $k$  est un mot à  $s$  lettres formé sur un alphabet à  $k$  lettres  $(l_1, \dots, l_k)$  tel que  $(l_1, \dots, l_k)$  soit consistant. Pour choisir un tel mot, il faut donc choisir  $k$  lettres parmi  $p_1, \dots, p_n$ , soit  $C_n^k$  choix possibles. Pour chaque lettre, il faut ensuite choisir entre la lettre et son contraire, soit  $2^k$  choix possibles. Puis choisir une séquence  $(n_1, \dots, n_k)$  telle que  $\sum_i n_i = s$ , soit  $C_{s-1}^{k-1}$  choix possibles. Le cardinal de l'ensemble des formules de type  $k$  est donc  $2^k C_n^k C_{s-1}^{k-1}$ . D'autre part le cardinal de  $\Phi_s$  est  $\sum_{i=1}^s C_{2n}^i C_{s-1}^i$  : c'est la somme des cardinaux de mots de longueur  $s$  comprenant exactement  $i$  lettres différentes sur un alphabet de cardinal  $2n$ . Ceci nous donne la formule annoncée. ■

Nous trouvons ainsi par complémentation  $\mu_\infty = 1 - \frac{\sum_{k=1}^s 2^k C_n^k C_{s-1}^{k-1}}{\sum_{i=1}^s C_{2n}^i C_{s-1}^i}$ .

## 7.2. Les variables globales

Comme nous l'avons montré dans la seconde partie, nous avons pour une stratégie  $(\nu_n, \nu_\lambda, \nu_{\bar{\lambda}})$

$$\begin{aligned} f_k &= \sum_{t=1}^{\lambda} \alpha^t \cdot \frac{1}{2^k} \left(1 - \frac{1}{2^k}\right)^{t-1} \\ &= \frac{\alpha}{2^k} \frac{1 - \alpha^\lambda \cdot \left(1 - \frac{1}{2^k}\right)^\lambda}{1 - \alpha \cdot \left(1 - \frac{1}{2^k}\right)} \end{aligned}$$

et:

$$\bar{f}_k = \sum_{t=1}^{\bar{\lambda}} \alpha^t \cdot \frac{1}{2^k} \cdot \left(1 - \frac{1}{2^k}\right)^{t-1}$$

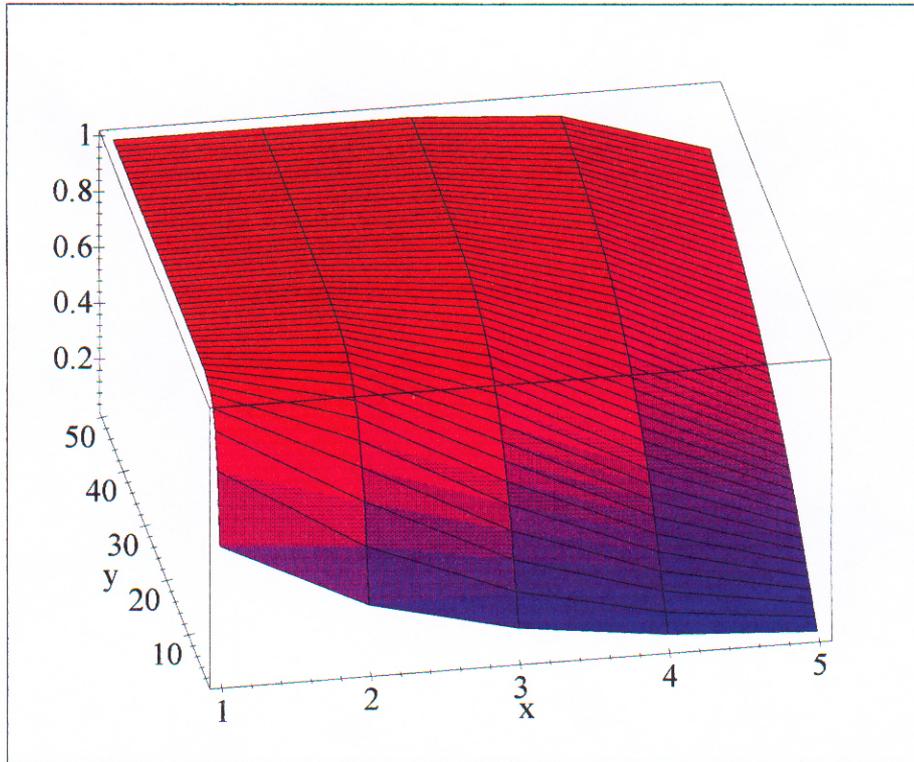
Pour se faire une idée, étant donné que  $\alpha \approx 1$ , nous avons:

$$f_k \approx 1 - \alpha^\lambda \cdot \left(1 - \frac{1}{2^k}\right)^\lambda$$

et

$$\bar{f}_k \simeq 1 - \alpha^{\bar{\lambda}} \cdot \left(1 - \frac{1}{2^k}\right)^{\bar{\lambda}}$$

Si nous traçons le graphe de cette approximation (que nous ne considérerons pas lors les calculs ultérieurs) nous pouvons nous représenter l'évolution des  $f_k$  et  $\bar{f}_k$  en fonction du temps de recherche moyen, pour par exemple,  $\alpha = 0,9995$  (ce que nous prendrons pour les simulations), qui correspond à une durée de vie moyenne de 2000 pas de temps:



Dépendance de  $f_k$  vis à vis du type (x) et du temps de recherche (y).

### 7.3. Evolution des proportions des différents types

**Cas général.** Nous allons, en vue d'une simulation, établir l'évolutions des différents domaines de  $\Phi$  en fonction de l'âge de la communauté  $\tau$ . Pour cela, nous allons considérer que les paramètres du modèle sont relativement constants sur un intervalle de temps  $\eta$ , pendant lequel les agents conservent la même stratégie. Nous rappelons que la distribution des théories dans  $\overline{CP}$  est représentée par  $(\bar{\mu}_k(\tau))$ , que celle de  $V$  est représentée par  $(\mu_k(\tau))$ , et celle de  $F$  par  $(\mu'_k(\tau))$ . Nous aurons aussi les dépendances  $\sigma(\tau)$ ,  $\theta(\tau)$ ,  $m(\tau)$ ,  $\lambda(\tau)$ ,  $\bar{\lambda}(\tau)$ ,  $\langle \nu \rangle(\tau)$ . Sur une période de durée  $\tau' - \tau = \eta$ , caractérisée par les temps d'arrêt  $(\nu_n, \nu_\lambda, \nu_{\bar{\lambda}})$  nous avons alors:

- modifications de  $\overline{CP}$  : En moyenne le nombre de processus de recherche menés durant cette période est:

$$\begin{aligned} nb &= \frac{N.\eta}{(1-\ell)\langle\nu_{\lambda}\rangle + \ell.\langle\nu_{\lambda}\rangle + \langle\nu_n\rangle} \\ &= |\Phi| \cdot \frac{r.\eta}{(1-\ell)\langle\nu_{\lambda}\rangle + \ell.\langle\nu_{\lambda}\rangle + \langle\nu_n\rangle} \end{aligned} \quad (7.1)$$

La variation du nombre de théories de type  $k$  dans  $\overline{CP}$  est alors:

$$\overline{\Delta}_k = -\overline{\mu}_k(\tau) \cdot (1-\ell) \cdot nb \cdot (1 - \overline{f}_k)$$

donc:

$$1 - \sigma_{\tau'} = \max\left((1 - \sigma_{\tau}) + \sum \frac{\overline{\Delta}_k}{|\Phi|}, 0\right)$$

et

$$\overline{\mu}_k(\tau') = \max\left(\overline{\mu}_k(\tau) \cdot \frac{1 - \sigma}{1 - \sigma'} + \frac{\overline{\Delta}_k}{|\Phi| \cdot (1 - \sigma')}, 0\right)$$

Toutes ces formules sont valables pour  $k = \infty$  avec  $\infty, f_{\infty} = 0$

Nous en déduisons  $\sigma_{\tau'}$ .

- modifications de  $V$  : celles-ci ont deux causes, l'arrivée de nouvelles formules venant de  $\overline{CP}$  qui ont échappé à la réfutation, l'élimination de formules réfutées.

Le nombre de formules de type  $k$  venant de  $\overline{CP}$  est :  $\max\left(\left(1 - \frac{r.(1-\ell).\langle T \rangle}{(1-\sigma).\langle \nu \rangle} \cdot f_k\right), 0\right) \cdot \overline{\Delta}_k$ ,  
 $\left(\frac{r.(1-\ell).\langle T \rangle}{(1-\sigma).\langle \nu \rangle} \cdot f_k\right)$  est la probabilité pour qu'une formule de type  $k$  soit réfutée immédiatement).

Le nombre de formules examinées auxquelles il est trouvé un contre exemple est :  $\mu'_k(\tau) \cdot \theta \cdot \ell \cdot nb \cdot f_k$ . Donc la variations du nombre de formules de type  $k$  dans  $V$  est :

$$\Delta_k = \max\left(\left(1 - \frac{r.(1-\ell).\langle T \rangle}{(1-\sigma).\langle \nu \rangle} \cdot f_k\right), 0\right) \cdot \overline{\Delta}_k - \mu_k(\tau) \cdot \theta \cdot \ell \cdot nb \cdot f_k$$

. Nous obtenons alors:

$$\theta_{\tau'} \cdot \sigma_{\tau'} = \min\left(1, \max\left(\theta_{\tau} \cdot \sigma + \sum \frac{\Delta_k}{|\Phi|}, 0\right)\right)$$

dont nous déduisons  $\theta_{\tau'}$  et

$$\mu_k(\tau') = \max\left(\mu_k(\tau) \cdot \frac{\theta_{\tau} \cdot \sigma_{\tau}}{\theta_{\tau'} \cdot \sigma_{\tau'}} + \frac{\overline{\Delta}_k}{|\Phi| \cdot \theta_{\tau'} \cdot \sigma_{\tau'}}, 0\right)$$

Nous obtenons ainsi les paramètres de la société pour la phase suivante.

**Phase initiale.** Pour la phase initiale, les calculs sont différents du fait qu'il n'y a pas encore de  $CP$  constitué. Ainsi,  $\sigma(0)$  est presque nul,  $(\mu_k(0))$ ,  $(\mu'_k(0))$  et  $\theta(0)$  dépendent des stratégies adoptées par les agents au cours de leurs recherches dans  $\overline{CP}$ . Nous allons voir comment nous pouvons calculer tous les paramètres du système en fonction de la distribution initiale  $(\bar{\mu}_k)$  des types dans la nature. Il nous faut ici trouver la distribution  $(\mu_k(0))$  à la formation de  $CP$  ainsi que le paramètre  $\theta(0)$ . Commençons par remarquer que la majeure partie des processus de recherche s'effectuant dans  $\overline{CP}$ , le temps de recherche moyen est en fait le temps de recherche moyen dans  $\overline{CP}$ . De plus, il y a peu d'exploration des formules. En effet, la proportion de  $CP$  est trop faible pour qu'il soit avantageux de rejeter une formule de  $\overline{CP}$  pour essayer d'en obtenir une de  $CP$  en un temps raisonnable. L'exploration ne peut se faire qu'en faveur des théories de  $\overline{CP}$ , c'est à dire après avoir tiré une théorie de  $CP$ , ce qui est très rare.

Ceci étant, nous avons une bonne approximation de  $\langle \nu \rangle(0)$  en considérant le temps de recherche dans  $\overline{CP}$  :

$$\langle \nu \rangle(0) = \langle \nu_{\overline{CP}} \rangle \quad (7.2)$$

Nous pouvons aussi calculer la probabilité  $pb(0)$  pour qu'il ait été trouvé un contre exemple à une formule donnée de  $\overline{CP}$  :

$$pb(0) = \min \left( 1, \frac{r \cdot (1 - \ell)}{(1 - \sigma)} \cdot \frac{\langle T \rangle}{\langle i\nu \rangle} \cdot \bar{f}_k \right)$$

Nous pouvons alors calculer la distribution  $(\mu_k(0))$ .  $\mu_k(0)$  est la proportion de théories de type  $k$  retenues parmi l'ensemble des publications acceptées. Un agent a la probabilité  $\bar{\mu}_k$  de tirer une formule de type  $k$  et la probabilité  $(1 - \bar{f}_k)$  de la proposer en fin de compte. Cette formule à ensuite la probabilité  $pb(0)$  d'être réfutée immédiatement. Il s'en suit que la distribution  $(\mu_k)$  est définie par:

$$\mu_k(0) = \frac{(1 - f_k) \cdot \mu_k \cdot pb(0)}{\sum_{k=1}^s (1 - f_k) \cdot \mu_k \cdot pb(0)}$$

Pour le paramètre  $\theta$ , nous pouvons dire qu'au tout début, les théories qui figurent dans  $F$ , sont celles qui ont été réfutés dès leur publication, puisque la société est trop jeune pour qu'il y ait des théories de  $CP$  réfutées a posteriori.  $\theta(0)$  coïncide donc avec la probabilité pour qu'une théorie soit réfutée lors de sa publication. Donc:

$$\theta(0) = pb(0)$$

Nous avons ainsi les paramètres initiaux de la communauté. L'étude précédente nous permet alors de voir comment ils évoluent avec le temps.

#### 7.4. Les résultats de la simulation

Nous disposons encore de trop peu de résultats pour pouvoir en faire un compte rendu précis. Nous les exposerons lors de la soutenance. Toutefois nous pouvons faire quelques

remarques relatives aux observations que nous avons recueillies. Elle ont été effectuées avec les coefficients suivants:

$$s = 5$$

$$n = 30$$

$\theta_k$	$1 - \frac{1}{2}$	$1 - \frac{1}{2^2}$	$1 - \frac{1}{2^3}$	$1 - \frac{1}{2^4}$	$1 - \frac{1}{2^5}$	1
$\bar{\mu}_k$	0	0	0.02	0.23	0.6	0.15

$\alpha = 0.9995$ , temps de vie de 2000 pas de temps.

$$r=0.0003$$

de telle sorte que  $\bar{p}$  initialement 0,6. Quand un agent tire une formule de  $\bar{p}$  celle-ci a déjà été examinée par un agent avec la probabilité 0,6.

- L'espérance de gain  $X$  décroît au cours de la simulation.
- Pour des valeurs de  $R$  faibles, de l'ordre de quelques unités, les agents préfèrent les formules de  $\overline{CP}$  à celles de  $CP$  au moins jusqu'à un âge de la communauté de quelques milliers de pas de temps. Au départ,  $\lambda$  et  $\bar{\lambda}$  valent 1, les agents proposent des formules sans vraiment les tester et ne sont pas très friands de vérification. Lorsque  $\sigma$  avoisine les 0.5, les comportements changent, les agents préfèrent toujours les formules de  $\overline{CP}$  à celles de  $CP$  mais lorsqu'ils entament une procédure de vérification, ils la mènent sur quelques dizaines de pas de temps. Il y a donc une phase de vérification des formules de  $CP$ .
- Pour des valeurs de  $R$  élevées, quelques dizaines d'unités, il y a changement de préférence au cours de l'évolution de  $\overline{CP}$  vers  $CP$  et la phase de vérification apparaît plus tôt avec des temps d'arrêts plus longs.

## Part IV

# Vers un modèle plus général

Nous allons proposer dans cette section des extensions du modèle logique précédent. La plupart seront sous forme d'idées et de conjectures. Parfois, nous donnerons un plan de démonstration.

### 8. L'émergence des paradigmes

On pourrait regretter dans le modèle précédent que les agents ne puissent pas confronter les théories qu'ils étudient aux théories déjà existantes, et utiliser ainsi leur culture scientifique afin de former un jugement a priori sur les formules qu'ils étudient. Comme le fait remarquer Popper, "*la connaissance ne saurait s'élaborer à partir de rien, ni procéder de la seule observation. Les progrès du savoir sont essentiellement la transformation d'un savoir antérieur.*" [8] Examinons donc les conséquences d'une telle structure sur l'ensemble des connaissances scientifiques. Nous considérerons désormais que la répartition des probabilités effectives  $\theta_k$  est *quelconque* mais qu'elle détermine un ensemble fini ordonné  $(\theta_1, \dots, \theta_n, \theta_\infty)$  en correspondance avec les types  $T_k$ . Nous pouvons donc avoir n'importe quelle distribution sur les  $(\theta_k)$ . En suivant le principe de maximisation des espérances de gain prôné par la théorie des jeux, nous pouvons tenter de définir un mécanisme de choix des théories qui rendrait compte de l'idée de culture scientifique et de tradition scientifique.

Partons de l'idée que l'importance d'une théorie qui a été publiée est due au fait que nous faisons des anticipations sur le nombre de test qu'elle a déjà passé et que nous donnons une certaine confiance à l'expérimentateur pour avoir mené ces tests de manière rigoureuse. Dans la logique de notre modèle, un agent peut donc donner une valeur a priori à une théorie s'il peut évaluer la manière dont celle-ci est corroborée par les théories déjà existantes. Regardons donc de quelle manière il serait possible d'évaluer la corroboration d'une théorie  $\varphi$  par une théorie  $\psi$ . *Nous supposons que les agents sont capables de comparer les extensions des différentes théories et de les intersecter.* Etant donné qu'une théorie ne prend de risques que sur ce qu'elle interdit, c'est l'ensemble  $\mathcal{E}_{\bar{\varphi}}$  qui va être intéressant pour déterminer l'indice de corroboration. D'autre part, la valeur d'une théorie  $\psi$  de  $CP$ , ne vient que du fait qu'elle a résisté à des tests, c'est à dire qu'elle indique que son extension  $\mathcal{E}_\psi$  contient des descriptions de phénomènes réels compatibles avec les tests qui lui ont été appliqué. C'est donc cet ensemble  $\mathcal{E}_\psi$  qui doit être pris en compte dans la corroboration de  $\varphi$  par  $\psi$ . Un agent cherche à connaître la probabilité effective de la théorie qu'il étudie, c'est à dire, la probabilité qu'elle a de résister à un test. Un agent qui tire une formule  $\varphi$  peut alors

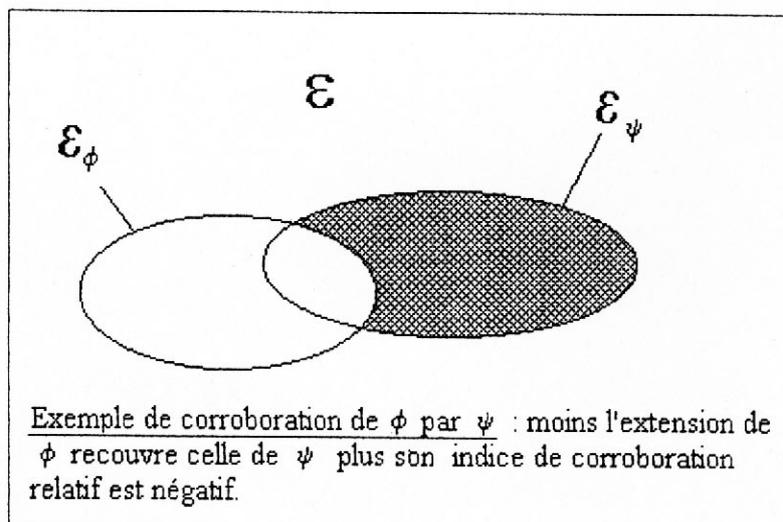


Figure 9:

comparer les extensions de  $\varphi$  et  $\psi$ , et en tirer la probabilité pour que les tests corroborants appliqués à  $\psi$  soient des contre-exemples pour  $\varphi$ . Pour ce faire, un agent doit évaluer le nombre  $\langle \nu_{\bar{\lambda}} \rangle$  de tests qu'en moyenne  $\psi$  a pu subir avant d'être publiée. Il doit d'autre part déterminer la manière dont elle a pu être vérifiée, ceci dépendant de son ancienneté dans  $CP$  et de la manière dont les agents vérifient les formules (i.e. de  $\langle \nu_{\lambda} \rangle$ ). Il obtient ainsi un évaluation du nombre  $N_t$  de tests qu'a pu passer  $\psi$  en fonction de son ancienneté  $t$ .  $N_t$  est de la forme  $N_t = \langle \nu_{\bar{\lambda}} \rangle + c(t) \cdot \langle \nu_{\lambda} \rangle$ . Il doit ensuite calculer la proportion  $\mathcal{E}_{\psi}$  de l'extension de  $\psi$  qui se trouve dans le complémentaire de  $\mathcal{E}_{\varphi}$  et en déduire le nombre potentiel  $N_t \cdot \frac{|\mathcal{E}_{\psi} \cap \mathcal{E}_{\bar{\varphi}}|}{|\mathcal{E}_{\psi}|}$  de contre-exemples pour  $\varphi$  qui auraient servi à corroborer  $\psi$ . Nous pouvons alors poser  $I[\varphi, \psi] = -N_t \cdot \frac{|\mathcal{E}_{\psi} \cap \mathcal{E}_{\bar{\varphi}}|}{|\mathcal{E}_{\psi}|}$ ,  $I$  est l'indice de corroboration de  $\varphi$  par  $\psi$  (cf. figure 9)

Nous obtenons ainsi un indice de corroboration a priori pour  $\varphi$  par les formules de  $CP$  :

$$I_{\varphi} = \sum_{\psi \in CP} I[\varphi, \psi]$$

Quelques remarques sur cet indice de corroboration a priori. Il n'est pas symétrique. Cette dissymétrie est une conséquence directe de la dissymétrie qu'établit Popper entre vérité et fausseté. Du fait qu'une théorie ne dit rien de ce qui est permis, qu'elle ne fait qu'interdire, elle ne peut en aucun cas corroborer positivement une théorie. L'indice de corroboration est donc toujours négatif. D'autre part, si l'on considère la théorie  $\Xi$  qui décrit le monde réel, nous voyons que pour une théorie  $\varphi$ , type  $k$ ,  $I[\varphi, \Xi] = -N_t \cdot (1 - \theta_k)$ . Ceci relie l'indice de corroboration aux  $\theta_k$ . Cet indice de corroboration peut être utilisé par les agents afin de juger de l'opportunité d'une recherche. En dessous d'un certain seuil de corroboration, la théorie ne serait pas examinée. Le calcul de cet indice pourrait être le fruit de l'expérience de l'agent et de sa culture scientifique. Plus un agent a d'expérience,

plus il est capable de comparer les conséquences d'une théorie avec celles déjà existantes ce qui lui permet de calculer l'indice de corroboration de manière plus fine. Il ne semble pas difficile de démontrer, mais nous le poserons comme simple *conjecture*, que l'utilisation de cet indice augmente strictement l'espérance de gain d'un agent. Un jeu qui prendrait en compte cet indice de corroboration devrait donner lieu à des systèmes dans lesquels nous assisterions à l'émergence de paradigmes. En effet, au début de l'exploration d'un champ de recherche,  $CP$  est négligeable et les agents cherchent un peu au hasard les différentes théories sans influence de ce qui a été fait précédemment. Petit à petit, les agents prennent en compte l'indice de corroboration afin d'obtenir des informations a priori sur les théories qu'ils examinent. Ils se fixent un indice de corroboration a priori en dessous duquel ils rejettent les théories. Ce seuil existe dès lors que l'indice de corroboration a priori est considéré comme pertinent. Du fait de la présence de  $\mathcal{E}^{réel}$  un déséquilibre se produit dans la répartition des extensions des théories acceptées et il se dessine un domaine au sein duquel il ne se forme plus de théorie. Ce déséquilibre ne peut que s'accroître avec le temps. Nous devrions donc voir apparaître des zones de recherche négligées par les agents. Le complémentaire de ces zones de recherche représenterait un paradigme, ou, au niveau du chercheur, une heuristique. (cf. figure 10). Il serait d'ailleurs intéressant de voir comment nous pourrions introduire cet indice de corroboration dans les règles de dérivation d'un système logique afin de construire une "logique de la découverte scientifique".

Donnons une idée de plan de démonstration pour montrer ce phénomène d'émergence:

- Nous pouvons assimiler le monde à la théorie  $\Xi$ . L'indice de corroboration  $I[\varphi, \Xi] = N_i \cdot \theta_k$  nous montre alors que plus une théorie a une extension proche de celle de  $\Xi$ , plus elle a de chances d'être acceptée. Ceci fait qu'il y a une concentration des extensions des théories acceptées autour de celle de  $\Xi$ , i.e. du monde réel.
- Un fois que le phénomène de concentration des extensions s'est développé, l'indice de corroboration a priori l'accroît pour donner naissance à un paradigme. Il faut alors montrer que le paradigme biaise le mécanisme d'acceptation des théories en déplaçant les extensions acceptées vers le cœur du paradigme.

Si l'on regarde l'effet de la présence d'un paradigme sur le développement des connaissances, nous constatons que son rôle est d'augmenter les proportions de bonnes théories dans l'ensemble des théories sur lesquelles un agent est susceptible d'effectuer des recherches. Son rôle est donc de changer la distribution des  $(\mu_k)$ . La simulation construite précédemment nous donne donc l'allure de ce qui se passe en présence d'un tel mécanisme.

D'un autre côté, si l'on permet à un contre exemple qui a été opposé à une théorie de  $CP$ , de réfuter toutes les théories avec lesquelles il est en contradiction, nous devrions voir apparaître, lorsqu'une nouvelle expérience arrive au centre de l'extension de plusieurs théories

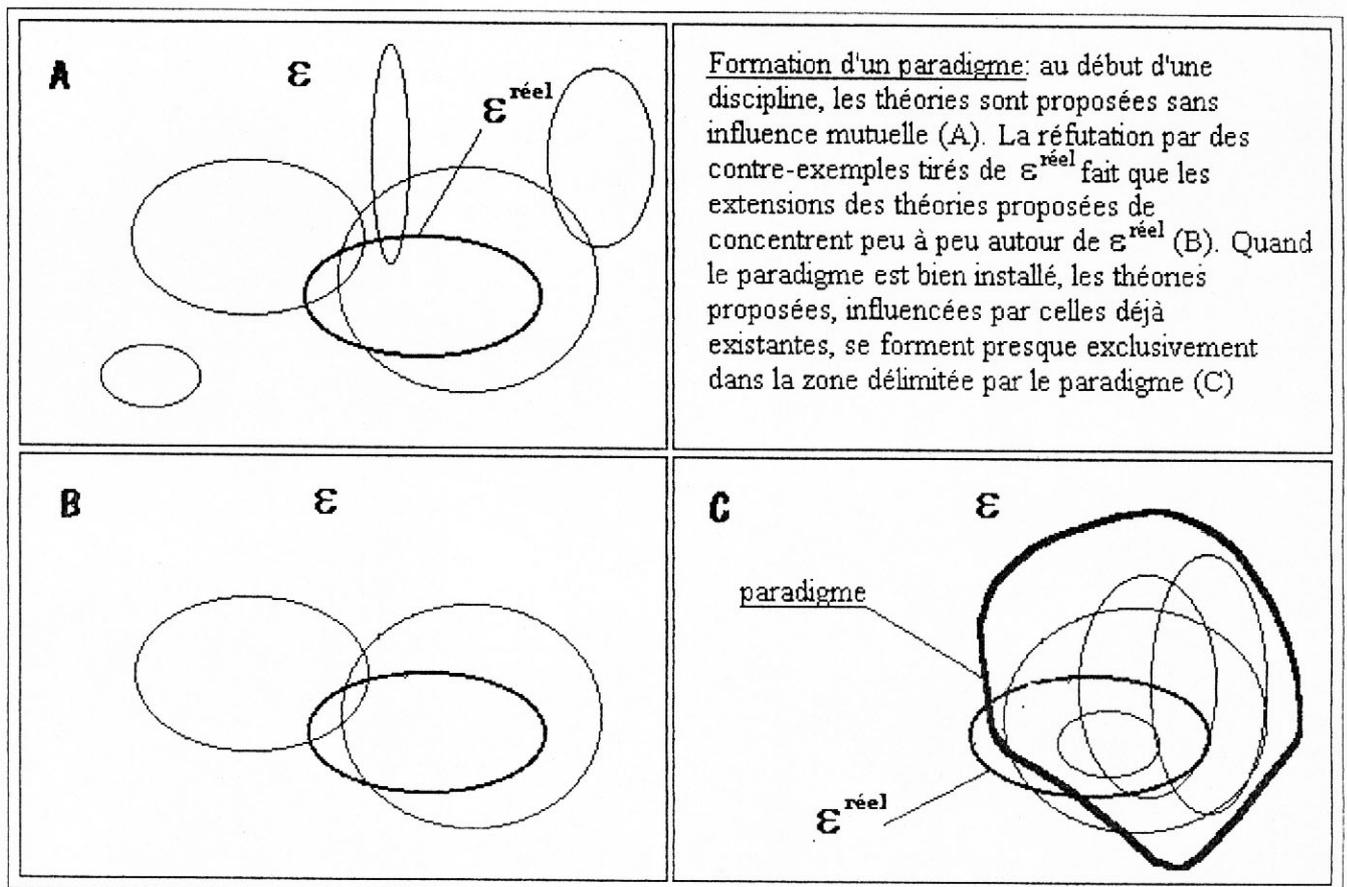


Figure 10:

emboîtées, des phénomènes semblables aux révolutions scientifiques dont parle Kuhn dans "La structure des révolutions scientifiques". A la suite d'un tel phénomène, le paradigme peut changer et de nouvelles directions de recherche peuvent apparaître. Si le paradigme est trop endommagé, il se peut aussi qu'il y ait une phase semblable à la phase initiale où les agents effectuent des recherches quasiment au hasard.

### 9. La coévolution de la théorie et de la technologie

Le paragraphe précédent amène naturellement une question: comment peuvent arriver des expériences susceptibles de provoquer une révolution scientifique? Mis à part les phénomènes extraordinaires qui auraient échappé à la plupart des scientifiques, il est très fréquent que des expériences cruciales pour un paradigme apparaissent après un saut technologique. De nouvelles mesures permettent d'observer des phénomènes nouveaux qui mettent en défaut les théories existantes. Comment rendre compte de cette dimension dans le modèle précédent? Pour cela, il nous faut regarder la forme des extensions des formules disjonctives en fonction des paramètres  $r$  et  $s$ .

Une formule est d'autant plus informative qu'elle comporte moins de variables propositionnelles. Par exemple  $\varphi = p_n$  interdit la moitié des événements, ceux pour lesquels  $p_n = 0$  alors que  $\psi = p_i \vee p_j \vee p_k$  interdit  $\frac{1}{8}$  des événements, ceux pour lesquels ( $p_i = 0$  et  $p_j = 0$  et  $p_k = 0$ ). Donc, plus une théorie comporte de variables propositionnelles, plus elle est sélective dans ce qu'elle interdit et plus elle permet de décrire le monde de manière plus fine.

La constante  $n$ , peut en revanche être interprétée comme un niveau d'avancée technologique de la société. En effet, plus  $n$  est grand, plus il y a de points dans l'ensemble des énoncés de base et donc plus il y a de mesures différentes possibles que l'on peut effectuer dans le monde réel. Donc plus  $n$  est grand, plus une société peut catégoriser les phénomènes naturels en les distinguant par des mesures fines. Nous pouvons à partir de là, proposer un type de communauté d'agent présentant une évolution tant sur le plan théorique que sur le plan technologique.

Soit  $n \in \mathbb{R}$ , considérons l'ensemble  $\varphi_n$  des formules logiques disjonctives sur  $p_1, \dots, p_n$  et l'ensemble  $\mathcal{E}_n$  des valeurs possibles de  $(p_1, \dots, p_n)$ . Nous définirons une  $k$ -formule  $\varphi_k$  à partir d'une formule  $\varphi \in \Phi_n$ , la formule obtenue en ne gardant que les  $k$  premières variables propositionnelles (par exemple, si  $\varphi = q_1 \vee q_2 \dots \vee q_{k-1}$ ,  $\varphi_2 = q_1 \vee q_2$ ). De même nous définirons une  $k$ -expérience  $e_k$  à partir d'une expérience  $e$ , le booleen obtenu en ne gardant que les  $k$  premières valeurs de  $e$  (par exemple, si  $e = 11001\dots 1$ ,  $e_4 = 1100$ ). Nous définirons alors l'action de  $\varphi_k$  sur  $e_r$  par  $\varphi_k(e_r) = \varphi_{k \wedge r}(e_{k \wedge r})$ . Les formules telles que  $\varphi_{k \wedge r} = \emptyset$ , si  $k$  et  $r$  sont les niveaux d'avancée théorique et technologique de la société, n'ont pas de valeur de vérité, elle peuvent toutefois engendrer comme le remarque Popper, un système métaphysique<sup>2</sup>. Nous avons alors un système qui peut subir des sauts théoriques

<sup>2</sup>Voir "La logique de la découverte scientifique", ch: "corroboration ou comment une théorie résiste à

aussi bien que des avancées technologiques. A niveau technologique donné, deux théories dont le niveau théorique excède le niveau technologique peuvent avoir les mêmes valeurs de vérité en tout point du monde réel et pourtant différer sensiblement sur la forme. Notamment, deux théories contradictoires pourront être également plausibles si elles diffèrent sur des prédictions qui ne seraient vérifiables qu'à un niveau technologique plus avancé. Opter pour l'une ou pour l'autre tiendra alors de la métaphysique. Cette contradiction sera une motivation pour améliorer les outils technologiques du moment afin de trouver de nouveaux contre-exemples (rappelons qu'un des résultats de la simulation est que l'espérance de gain  $X(n, \lambda, \bar{\lambda})$  diminue avec l'âge du domaine de recherche au fur et à mesure que  $\overline{CP}$  s'épuise et que  $F$  s'enrichit de toutes les théories qui ne sont de type fini ( $\notin T_\infty$ ). Elargir  $\mathcal{E}$  et donc  $\Phi$  serait un bon moyen d'augmenter l'espérance de gain). De même, à la suite d'un saut technologique, certaines théories proposées pourront se révéler fausses à cause de résultats expérimentaux plus précis, ce qui poussera rejeter des théories anciennement acceptées pour en adopter de nouvelles. Il faut remarquer aussi que dans ce modèle, il n'y a plus obligatoirement de théorie générale finie du monde  $\Xi$ , celle-ci peut comporter un nombre infini de clauses, elles-mêmes de taille infinie. Pourtant à technologie et à niveau théorique donnés  $r$  et  $s$ ,  $\mathcal{E}(r, s) = \{e_r \text{ tel que } e \in \mathcal{E}\}$  est fini ainsi que  $\Phi(r, s) = \{\varphi_r \text{ tel que } (\varphi \in \Phi) \wedge (\varphi \text{ de taille inférieur à } s)\}$ . Le monde perceptible peut alors être décrit par une formule générale  $\Xi_{r,s}$ . Rappelons que les agents cherchent à constituer un  $CP$  qui approche le mieux possible la grande théorie du monde  $\Xi$  (qu'ils ne peuvent en principe jamais atteindre). La conception  $\chi_{r,s}$  que se fait la société du monde, représentée par l'ensemble des connaissances de  $CP$ ,  $\chi_{r,s} = \bigwedge_{\varphi \in CP} \varphi$ , sera alors d'autant plus précise que  $s$  et  $n$  seront grands. Nous voyons donc sous ce modèle guidé par le principe de réfutation, théorie et technologie évoluent parallèlement pour amener la société vers une description du monde de plus en plus précise. Il faudrait cependant, pour que le modèle soit complet, introduire un mécanisme qui permette d'éliminer les théories les moins informatives au profit des théories les plus informatives.

## 10. Conclusion

*"Il n'y a qu'une seule façon de présenter des arguments en faveur de mes propositions. C'est d'analyser leurs conséquences logiques, mettre en évidence leur fécondité, leur pouvoir d'élucider des problèmes de la théorie de la connaissance."*

Certains antagonistes de Popper ont reproché à sa thèse de ne pas avoir la forme qu'elle-même préconisait pour les théories scientifiques. Pourtant, si l'on regarde de près la forme de la théorie de la réfutation, nous pouvons voir que celle-ci décrit les processus de développement de la connaissance de manière assez précise et assez pure. Les contraintes qu'elle impose au développement des connaissances scientifiques devraient en fait, avoir des conséquences observables. Ce sont ces conséquences que nous nous sommes proposés de mettre en évidence par l'intermédiaire de la modélisation.

Nous avons fait attention à poser des principes aussi simples que possibles et en accord avec le sens commun, afin d'obtenir des résultats qui puissent, dans une certaine mesure, être confrontés à des phénomènes observables. Nous avons ainsi explicité une famille de temps d'arrêt susceptibles de guider la recherche scientifique. Ceux-ci sont assez intuitifs puisqu'ils correspondent à la donnée d'une date butoire au delà de laquelle toute recherche doit être interrompue quels que soient les résultats. Nous avons d'autre part explicité la dépendance de la qualité des théories reconnues par rapport à ces temps d'arrêt. Enfin les résultats des simulations semblent montrer, ce qu'il faudrait confirmer par un ensemble plus complet d'expériences, que la thèse de Popper a une conséquence importante sur le développement des connaissances scientifiques et les comportements des agents. Nous avons en effet constaté que le développement d'une communauté s'effectuait en plusieurs phases bien distinctes. Au début, il y a une phase d'exploration de l'espace des formules durant laquelle les agents publient beaucoup sans tellement s'assurer de la validité de leurs théories et d'un autre côté, s'intéressent peu à la vérification des théories publiées par leurs collègues. Puis peu à peu, les temps de recherche augmentent et les agents s'attachent de plus en plus à la vérification des théories publiées jusqu'à ce que les temps d'arrêt se stabilisent. Il y a alors une phase de consolidation de l'ensemble des connaissances.

Nous sommes tout à fait conscients du fait que le modèle proposé ici est très simplifié et peut sembler bien loin de la réalité. Il représente surtout un schéma possible pour penser la dynamique des sciences. Cependant, la voie de la modélisation semble fructueuse dans l'étude de la thèse de Popper du moment qu'elle n'a pas son emploi habituel. Elle ne doit pas être utilisée pour prédire ou expliquer certains phénomènes, mais bien comme un outil pour forger des tests, précisément au sens de Popper. Celle-ci nous permet de voir comment des règles de découverte au niveau de l'individu peuvent influencer la dynamique de com-

portements collectifs qui eux, sont observables. Des études plus poussées dans ce sens, ainsi que le développement des idées que nous avons avancées pour relier les thèses poppériennes aux mécanismes de formation des paradigmes, devraient permettre à la thèse de Popper de satisfaire à ses propres critères, en permettant de comparer le type de développement de la connaissance qu'elle imposerait à celui que nous avons pu observer au cours de l'Histoire.