

BULLETIN N° 150
ACADÉMIE EUROPEENNE
INTERDISCIPLINAIRE
DES SCIENCES



Séance du mardi 14 décembre 2010:

Présentation par Daniel KAYSER de l'Université Paris XIII de :
"Quelques réflexions sur le traitement automatique du langage naturel".

Prochaine séance mardi 11 janvier 2011 de 18h à 19h30 précises au
« Carré des Sciences » du Ministère de l'Enseignement Supérieur et de la Recherche
25, rue de la Montagne Sainte Geneviève* Paris 5^{ème} (Métro Maubert-Mutualité)

AMPHI A Rez de Chaussée haut

Présentation par le Pr SIBERTIN-BLANC de L'Université de Toulouse 1 de :
"des résultats de simulations à des connaissance en sciences sociales".

*une pièce d'identité sera demandée à l'entrée

ACADEMIE EUROPEENNE INTERDISCIPLINAIRE DES SCIENCES

FONDATION DE LA MAISON DES SCIENCES DE L'HOMME

PRESIDENT : Pr Victor MASTRANGELO
PRESIDENT Sortant: Michel GONDRAN
SECRETAIRE GENERAL : Irène HERPE-LITWIN
TRESORIER GENERAL : Bruno BLONDEL
MEMBRES DU CA Patrice CROSSA-RAYNAUD, Claude ELBAZ

PRESIDENT FONDATEUR : Dr. Lucien LEVY (†)
PRESIDENT D'HONNEUR : Gilbert BELAUBRE
SECRETAIRE GENERAL D'HONNEUR : Pr. P. LIACOPOULOS (†)

CONSEILLERS SCIENTIFIQUES :
SCIENCES DE LA MATIERE : Pr. Gilles COHEN-TANNOUDJI
SCIENCES DE LA VIE ET BIOTECHNIQUES : Pr François BEGON

SECTION DE NICE :
PRESIDENT : Doyen René DARS

SECTION DE NANCY :
PRESIDENT : Pr Pierre NABET

Décembre 2010

N°150

TABLE DES MATIERES

- P. 03 Compte-rendu de la séance du mardi 14 décembre 2010
- P. 05 Compte-rendu de Section Nice Côte d'Azur du 18/11/2010
- P. 08 Annonces
- P. 09 Documents

Prochaine séance: mardi 11 janvier 2011 de 18h à 19h30 précises au
« Carré des Sciences » du Ministère de l'Enseignement Supérieur et de la Recherche
25, rue de la Montagne Sainte Geneviève Paris 5^{ème}

AMPHI A rez-de-chaussée haut

Présentation par le Pr SIBERTIN-BLANC de l'Université de Toulouse1de :
"des résultats de simulations à des connaissance en sciences sociales"

**ACADEMIE EUROPEENNE INTERDISCIPLINAIRE DES
SCIENCES**

Maison des Sciences de l'Homme, Paris.

**Séance du
Mardi 14 décembre 2010**

Maison des Sciences de l'Homme, salle 214, à 18 h.

La séance est ouverte à 18 h. 00 sous la Présidence de Victor MASTRANGELO et en la présence de nos collègues Gilbert BELAUBRE, Gilles COHEN-TANNOUDJI, Claude ELBAZ , Michel GONDRAN, Walter GONZALEZ , Irène HERPE-LITWIN, Gérard LEVY, Jacques LEVY , Pierre MARCHAIS, Pierre PESQUIES .

Etaient excusés François BEGON, Bruno BLONDEL, Alain CARDON, Brigitte DEBUIRE, , Françoise DUTHEIL, Jean -Pierre FRANCOISE, Marie-Louise LABAT, Saadi LAHLOU, Alain STAHL.

D) Question diverses

- A) Notre Président Victor MASTRANGELO nous informe de l'obtention d'un accord avec l'Université Paris VII –Diderot pour héberger notre futur colloque les 28 et 29 novembre 2011
- B) La MSH étant indisponible, il nous faudra trouver un nouveau lieu de tenue de nos séances notamment pour janvier et février 2011.
- C) En vue d'une diminution de nos dépenses, notre Président d'honneur suggère la suppression de la diffusion du bulletin papier. Seul subsistera la version électronique du bulletin. En revanche, on pourrait concevoir une impression annuelle des annales des travaux de l'AEIS.
- D) Notre Collègue Gilles COHEN-TANNOUDJI insiste alors sur un archivage de ces travaux sur le site de l'AEIS.

II) **Présentation par le Pr Daniel KAYSER de l'Université Paris XIII de :** **« Quelques réflexions sur le traitement automatique du langage naturel »**

Daniel KAYSER exerce ses fonctions de professeur à l'Institut Galilée de l'Université Paris 13 - située 99 avenue Jean-Baptiste Clément 93340 VILLETANEUSE . L'Institut Galilée se consacre à l'enseignement , à la recherche en mathématiques, en physique, en chimie et en informatique, depuis 35 ans.

Il comporte des formations diverses allant de la préparation à la licence jusqu'à une école doctorale et abrite en son sein une école d'ingénieurs.

Au niveau de la recherche en informatique, il comporte un laboratoire , le LIPN (Laboratoire d'Informatique de l'université Paris Nord) très impliqué notamment dans les méthodes de fouilles de données (data mining) qui ont donné lieu à un colloque les 29 et 30 juin 2010 sur « l'apprentissage artificiel et la fouille de données ».

Daniel KAYSER est particulièrement motivé par la compréhension et la modélisation du raisonnement sémantique des Langues Naturelles, d'où son exposé.

Son exposé comportera donc une réflexion sur les modèles en général et plus particulièrement ceux qui sont sémantiques avec leurs aspects multiples, lexicaux, grammaticaux...

Voici un résumé de sa conférence que l'on pourra approfondir en consultant un tableau de ses diapositives dans les documents page 10:

Résumé de l'exposé du Pr Daniel Kayser

Les recherches du Professeur Kayser concernent les méthodes de représentation des connaissances adaptées à la compréhension du langage naturel ; elles se situent dans le cadre de l'Intelligence Artificielle et des Sciences Cognitives.

Daniel Kayser, dans son exposé intitulé « Le Traitement Automatique du Langage Naturel », nous a proposé le plan suivant : réflexions sur les modèles ; les modèles sémantiques : lexicaux, grammaticaux. Il argumente d'abord sur l'utilité d'un modèle à décrire un phénomène (de façon compacte) ; à accomplir certaines tâches (le modèle peut dépendre de la tâche elle-même). Quand le phénomène est complexe, on peut adopter une attitude cartésienne qui consiste à découper le phénomène en sous-problèmes ou bien adopter une attitude systémique mais avec le risque d'un modèle type « boîte noire » et enfin apparaît le problème de la plausibilité cognitive. Ensuite, il énonce les différents modèles possibles plus ou moins autonomes de chaque sous-domaine : lemmatiseurs, étiqueteurs, analyseurs syntaxiques et différents modèles sémantiques.

Il présente ensuite les modèles sémantiques pour lesquels il existe trois métaphores : linguistique, chimique (combien d'atomes de sens ?...), spatiale (qu'est-ce qu'une distance sémantique ? ..). Ces modèles possèdent trois niveaux de description : lexical, grammatical et textuel. Après avoir illustré ces niveaux par de nombreux exemples, il introduit les notions d'opérateurs de Prédicat, de Fonction, d'Ajout....avec de nombreux exemples. Il termine son exposé par l'analyse de l' « analogie synchronie-diachronie » utilisée dans l'analyse des textes (Sémantique).

 Après ce riche exposé et de nombreuses questions, la séance prend fin.

Bien amicalement à vous,

Comptes-rendus de la section Nice-Côte d'Azur

La science consiste à regrouper des faits
afin de pouvoir en tirer des lois ou
conclusions générales.
Charles Darwin.

Compte rendu de la séance du 18 novembre 2010 (142^{ème} séance)

Présents :

Patrice Crossa-Raynaud, Guy Darcourt, René Dars, Jean-Pierre Delmont, Jacques Lebraty, Maurice Papo.

Excusés :

Jean Aubouin, Richard Beaud, René Blanchet, Sonia Chakhoff, François Cuzin,
Jean-Paul Goux, Yves Ignazi.

1- Approbation du compte rendu de la 141^{ème} séance.

Le compte rendu est approuvé à l'unanimité des présents.

2- Conférences au Mamac.

Deux conférences sur le thème de la diversité ont été données au MAMAC :

- ▶ 26 octobre : P. Crossa-Raynaud : *La diversité dans la domestication des espèces,*
- ▶ 15 novembre : Professeur Jean-Pierre Delmont : *La diversité en médecine.*

devant un public restreint.

Cela était prévisible puisqu'il s'agit d'une nouveauté à Nice dans ce lieu.

Nous avons pourtant fait éditer des affiches annonçant ces conférences dont la deuxième a fait l'objet d'une présentation dans Nice-Matin.

Une discussion s'est engagée entre nous pour rechercher les moyens d'attirer un public plus nombreux.

Ces conférences vont être publiées sur le site de la Mairie de Nice, à une rubrique qui nous est réservée :

<http://etudiants.nice.fr/culture-scientifique,140.html>

Guy Darcourt a rencontré M. Stéphane Dupont qui lui propose de mettre en place des liens informatiques permettant au public de retrouver notre site plus facilement. Notre confrère Maurice Papo accepte d'étudier ce problème avec Guy Darcourt.

Notre confrère Pierre Couillet a accepté de faire connaître nos conférences au public des Lundis de la connaissance.

Programme du 1^{er} trimestre 2011

Annances

I) Claudine COHEN de l'EHESS nous fait part de son séminaire :

BIOLOGIE ET SOCIETE

Séminaire 2010-2011

sous la direction de

Henri ATLAN et Claudine COHEN

MODELISER LA COMPLEXITE EN BIOLOGIE

Lundi 13 décembre 2010

M. Henri ATLAN (EHESS)

« Y a-t-il une science de la complexité ? »

Lundi 10 janvier 2011 de 13 h à 15 h.

Mme Claudine COHEN (EHESS)

« Modèles et récits de l'évolution du vivant »

Lundi 13 février 2007 de 13 h à 15 h

M. Benoît RITTAUD (Université Paris XIII)

« La modélisation des changements climatiques :
Réflexions sur quelques aspects des débats actuels »

Lundi 13 mars 2011 de 13 h à 15 h

M. Vincent COURTILLOT (Paris VII-Diderot)

« Le rôle des extinctions en masse dans l'évolution des espèces: catastrophes et émergences »

Lundi avril 2011 de 13 h à 15 h

M. Jacques DEMONGEOT (CNRS)

« La robustesse des modèles pour les réseaux de régulation biologique »

Lundi 9 mai 2011 de 13 h à 15 h

Mme Soraya de CHADAREVIAN (UCLA)

« Biologie et société :
Perspectives historiques et philosophiques »

***Toutes les séances du Séminaire auront lieu de 13 h à 15h,
A l'EHESS, Salle M. et D. Lombard, 96 bd Raspail, 75006 Paris -***

Les exposés dureront une heure et quart environ, afin de laisser une large place à la discussion.

Documents

Pour approfondir le résumé de la présentation du Pr Daniel KAYSER sur "*Quelques réflexions sur le traitement automatique du langage naturel*", nous vous proposons u tableau de ses diapositives :

P. 10 quelques réflexions sur le traitement automatique du langage »

Pour préparer la conférence du Pr SIBERTIN –BLANC, celui-ci nous a communiqué deux articles :

p.21 : « the validity of simulation results” parue dans MAPS2 conférence d’avril 2010

P. 30: « Modèles formels pour reproduction des simulations à base d'agents » paru dans *Nouvelles perspectives en Sciences Sociale* 5, 2, p.135 - 149, 2010

Quelques réflexions sur le Traitement Automatique du Langage Naturel

*Daniel Kayser
L.I.P.N. Institut Galilée
Université Paris-Nord*

PLAN

- Réflexion sur les modèles
- Modèles sémantiques
 - Lexicaux
 - Grammaticaux
- Conclusion

À quoi sert un modèle ?

À décrire un phénomène (de façon compacte)

- Dans le cas du langage, le phénomène est-il
 - L'existence attestée d'une liste d'énoncés ?
 - La correspondance entre ces énoncés et des situations ?

À accomplir certaines tâches (mieux que sur l'objet lui-même)

- Le modèle peut alors dépendre de la tâche
 - Réponse aux questions
 - Traduction
 - Synthèse de documents ...

Quand le phénomène est complexe...

- Attitude cartésienne: découper en sous-problèmes
- Plus adaptée aux sciences «exactes» qu'aux sciences humaines
- Dans le cas du langage, le découpage : morphologie, lexicale, syntaxe, sémantique, pragmatique est-il satisfaisant ?
 - Attitude systémique (mais risque d'un modèle «boîte noire»)
 - Problème de la plausibilité cognitive

Simplification – Idéalisation

- Locuteur natif idéalisé
 - Accord inter-subjectif et même intra-subjectif généralement plus faible qu'attendu
 - Peu de prise en compte de la netteté du jugement
 - * Après une journée de travail, il a trouvé la résolution du problème (Pollock, 1990 p.50)
 - * Pierre semble que admire Paul (ibid. p.129)
- Risque de renvoyer les problèmes difficiles en fin de chaîne

Quels modèles ?

- Modèles ± autonomes de chaque sous-domaine
 - Lemmatiseurs
 - Étiqueteurs
 - Analyseurs syntaxiques
- structure complète

- «shallow parsers»
- Différents modèles sémantiques

Les frontières existent-elles ?

- **oui**, à gros grains
 - La conjugaison, le pluriel sont des phénomènes morphologiques
 - L'ordre des mots, l'accord sont des phénomènes syntaxiques
 - La référence, la compositionnalité sont des phénomènes sémantiques
 - La prise en compte du contexte d'énonciation est un phénomène pragmatique
- **Non**, à grain fin
Certains suffixes (-able, -iser, -isation...) à la frontière entre morphologie et lexique
L'acceptabilité d'une phrase à la frontière entre syntaxe, sémantique et pragmatique
 - La frontière entre le linguistique et l'extra-linguistique est une fiction !!

Modèles sémantiques

• 3 métaphores

Linguistique (Fodor, the Language of Thought, 1975)

Chimique

- Combien d'atomes de sens ?
- Règles de combinaison moléculaire ?

Spatiale

- Qu'est-ce qu'une distance sémantique ?
- Espace à combien de dimensions ?
- Un sens = un point, un champ scalaire, une forme ?

• 3 niveaux

- Lexical
- Combien de sens possède un mot ?
- Comment représenter un sens ?
- **Grammatical**
- Connaissant le sens des mots, quel est le sens de la phrase ?
- **Textuel:**
- qu'est-ce que la cohérence discursive ? Acceptabilité de séquences
 - de phrases
 - de paragraphes

Dialogues bien et mal construits

Sémantique lexicale traditionnelle

- Les mots dénotent des objets, des classes d'objets, ou des situations, états, processus, ... impliquant des objets d'un univers
- Cet univers (monde réel ou univers du discours) préexiste à l'interprétation
- cf. Bible : Dieu présente les animaux à Adam, et celui-ci leur donne un nom.

Sémantique lexicale les dictionnaires

- Aident les humains pour certaines tâches:

Orthographe

Mots rares : définition

Distinctions sémantiques fines

- Pas ce qui est déjà connu par tout utilisateur
- Présentent une «liste de sens»

Ce qu'est un sens n'est pas précisément défini

En fait, ≈ liste d'emplois

Exemple «**LAISSER**» (TLFI)

- Laisser-1

- Faire rester

- Quelqu'un, quelque chose laisse qq'un/qqch dans un lieu

Quitter momentanément

Quitter délibérément et définitivement

« Quelqu'un laisse quelque chose

« Quelqu'un laisse quelqu'un

Mettre en dépôt, en garde

Oublier

Perdre quelque chose qui nous appartient en propre

Faire subsister après soi

Faire rester après son passage

Abandonner après sa mort

- Quelqu'un, quelque chose laisse qq'un/qqch dans un état

Faire rester (le sujet est une personne)

(le sujet désigne une chose, l'objet désigne une personne)

- Faire garder

- Qq'un/qqch laisse qq'un/qqch à/pour qq'un/qqch

Remettre, confier

Céder

Garder, réserver

Accorder

Transmettre par succession

- Qq'un/qqch laisse qqch à faire à qq'un

Qq'un laisse qqch à (infinitif) à qq'un

Qq'un laisse (à qq'un) à penser/juger ...

Qq'un/qqch laisse (à qq'un) à (infinitif)

- Laisser-2

- Qq'un/qqch laisse qq'un/qqch (infinitif)

- Permettre de, ne pas empêcher

L'infinitif est intransitif ou pronominal

L'infinitif est transitif

Qq'un/qqch se laisse (infinitif): consentir à, céder à

«se» est sujet de l'infinitif

«se» est complément d'objet de l'infinitif

«se» est complément d'objet indirect de l'infinitif

Ne pas laisser de

Sémantique grammaticale compositionnelle

- Phrase → GN GV
- GN → Article indéfini Nom
- GV → Verbe transitif GN

Règles de construction associées:

GN → Article indéfini Nom

- prendre une nouvelle variable v
- trouver dans le lexique le prédicat N associé à Nom
 - fabriquer $\exists v N(v)$

GV → Verbe transitif GN

- trouver dans le lexique le prédicat V associé à Verbe transitif
- prendre la formule f associée à GN; soit v son 2èmesymbole
- ajouter à f : $\& V(_, v)$

Phrase → GN GV

- prendre la formule f associée à GN; soit v son 2èmesymbole
- prendre la formule g associée à GV; mettre v à la place de $_$
- fabriquer $f \& g$

Exemple: «un chien regarde un chat»

GN → Article indéfini Nom

- prendre une nouvelle variable v
- trouver dans le lexique le prédicat N associé à Nom
- fabriquer $\exists v N(v)$

GN $\exists x$ CHIEN(x)

GN $\exists y$ CHAT(y)

GV → Verbe transitif GN

- trouver dans le lexique le prédicat V associé à Verbe transitif
- prendre la formule f associée à GN; soit v son 2èmesymbole
- ajouter à f : $\& V(_, v)$

GV $\exists y$ CHAT(y) & REGARDER($_, y$)

• Phrase → GN GV

- prendre la formule f associée à GN; soit v son 2èmesymbole
- prendre la formule g associée à GV; mettre v à la place de $_$
- fabriquer $f \& g$

Phrase $\exists x$ CHIEN(x) & $\exists y$ CHAT(y) & REGARDER(x, y)

Sémantique lexicale: une autre modélisation

- Le principal besoin n'est pas la **définition** des mots, mais n'est pas la **définition** des mots, mais l'accès aux **inférences** qu'ils autorisent

il y a eu un problème de dysfonctionnement des feux de signalisation. Apparemment, le feu était au vert lorsque l'autre véhicule a poursuivi sa route. Je ne le conteste pas. Mais, de mon côté, le feu était bien vert.

«FEU» (TLFI)

I A •1. Phénomène consistant en un dégagement de chaleur et de lumière produit par la combustion vive d'un corps

Nombreux usages spéciaux, sens figurés, proverbes

- 2. Phénomène physique assimilé à une combustion vive
 - a. suivi d'un déterminant (*feu des volcans, feu follet*)
 - b. subst. + de feu (*langues de feu, ceinture de feu*)

- B Source de chaleur à fonction utilitaire ou symbolique

- 1. utilisée pour chauffer ou cuire

- a. feu + matière (*feu de bois*)

« α. Spéc. (*cuisinière 4 feux*)

« β loc. fig. (*feu de paille*)

- b. feu + déterm. spécifiant la fonction du feu (*feu de camp, feu de joie*)

II. Le feu dénote un des aspects du phénomène de combustion

- A. Aspect lumineux

- 1. Lumière produite par qqch qui brûle

- 2. (par extension) lumière

Militaire (*extinction des feux, couvre-feu*)

Marine, aviation, chemin de fer

Signalisation automobile

« Dispositif lumineux servant à signaler un véhicule automobile (*feux de stationnement, clignotants*)

« Dispositif lumineux dont est muni un véhicule pour éclairer la route (*feux de croisement*)

Circulation (*feux tricolores*) + sens figurés

Théâtre

Poésie

- 2. dans un syntagme prép. déterminant un type d'objet

a. usage domestique (*chambre à feu*)

b. usage médical (*pointes de feu*)

c. technologie (*pompe à feu, pierre à feu*)

d. sylviculture (*bois de feu*)

- 3. (par métonymie) **Maison habitée**

- C. Destruction de qqch par combustion vive (*mettre le feu*)

• Emplois spéciaux (*soldats du feu, feu de cheminée*)

• loc. fig. (*à feu et à sang*)

B. (métaphoriquement) un des aspects de la combustion, sans combustion réelle

- 1. -a. Chaleur vive

« Locutions (*eau de feu*)

« Emplois spéciaux (*feu du rasoir*)

- b. Domaine des pulsions et des affects
 - 2. Couleur \pm vive, comme celle du feu
 - III. Spécialement
 - A. Déflagration de matières fulminantes
 - B. (par métonymie) Affrontement armé
 - 1. Tir
 - Spéc. *feu roulant*
 - Loc. fig. *prendre entre deux feux*
 - 2. (généralement précédé d'un verbe/subst. indiquant le mouvement ou la situation dans l'espace)
- Combat

Le principal besoin: **feu de circulation**

- Situé avant certains carrefours fréquentés, pour éviter les conflits d'accès
 - Un feu par voie arrivant au carrefour. L'ensemble des feux d'un même carrefour peut être appelé «*le feu*»
 - Couleur changeant avec la phase :
Vert -> Orange -> Rouge -> Vert -> ...
- Opposition de phase entre les feux de voies orthogonales
Valeur déontique des couleurs

- Les accès aux **inférences** s'obtiennent par application (éventuellement itérée) d'un petit nombre d'**opérateurs**
 - Le dictionnaire donne des **points de départ**
- À partir desquels on applique ces opérateurs («**adaptation sémantique**»)
Le choix des points de départ n'est pas un problème critique
L'applicabilité des opérateurs peut parfois se déduire par héritage

Exemple : «**couleur**»

(i) *Cet homme est anglais*

(ii) *Les anglais sont des européens*

(iii) *Cet homme est un européen*

(1) Cette fleur est rouge

(2) Le rouge est une couleur

(3) Cette fleur est une couleur

(1) signifie en fait La couleur de cette fleur est rouge et «couleur» joue le rôle de fonction **couleur(fl#1)** = rouge

(2) Utilise «couleur» comme une prédication **couleur(rouge)**

couleur est une fonction, **couleur** est un prédicat.

- (autre solution: (**rouge(fl#1)**) \wedge **couleur(rouge)** avec **rouge** prédicat et **rouge**)

Ce problème est très général!!

Jean est le père de Sophie; c'est un père heureux!

- père: fonction, père: prédicat

- Le Havre est un port actif; Le port du Havre est actif

- port: prédicat, port: fonction

HPLIPN est une imprimante; c'est une des imprimantes du réseau B300

- imprimante: prédicat unaire; imprimante: prédicat binaire

- $\sin(x)$ est une fonction des réels sur les réels; $\sin(x)$ est un réel

- $\sin(x)$: nom d'une fonction; $\sin(x)$: valeur de cette fonction

[Montagne]the temperature is 90°F; the temperature is rising

- temperature: valeur de fonction; temperature: nom de fonction

Opérateurs

- *couleur* n'est pas interprétable
- Il existe un opérateur PRÉDICAT qui appliqué à *couleur* donne un prédicat unaire
 - Le rouge est une couleur \rightarrow PRÉDICAT(*couleur*) (rouge)
- Il existe un opérateur FONCTION qui appliqué à *couleur* donne une fonction unaire
 - Cette fleur est rouge \rightarrow =(rouge, FONCTION(*couleur*) (fl#1))
- Les autres exemples se traitent de façon semblable

Exemple : «**PRIX**»

- *prix* se comporte à première vue de la même façon
 - 15 € est un prix [prédicat]
 - 15 € est le prix de ce livre [fonction]
- Mais plusieurs facteurs montrent que la sémantique de *prix* est plus complexe
 - On peut parler du prix d'un objet existant en plusieurs exemplaires
 - On peut parler du prix d'une collection d'objets
 - On peut parler du prix d'entités massives
 - Le prix peut dépendre du temps et du lieu

Le prix de ce livre est 15 €

- Co-présence de deux entités :
 - *ce livre*: objet physique que je montre en ce moment
 - *ce livre*: classe d'objets ayant mêmes caractéristiques
- À partir de *x*, on construit CLASSE(*x*)
- On obtient =(15 €, FONCTION(*prix*) (CLASSE(*x*)))
- Une règle par défaut indique que

- si 15 € est le prix de la classe,
- c'est aussi le prix d'un élément de la classe,
- sauf exceptions (livre écorné, en exposition, ...)

L'opérateur AJOUT

- Si f est une fonction n -aire, $AJOUT(paramètre) f$ est une fonction $(n+1)$ -aire dont le $(n+1)$ ème argument est de type *paramètre*.
- *la baisse de 4,8% des prix des produits frais*
- $\exists p, t_0, t_1 = (p, AJOUT(temps)((FONCTION(prix)$
 (Collection-pondérée(x, AJOUT(temps)(P, t1))), t1)) \wedge $(*(p, 1, 048), AJOUT(temps)((FONCTION(prix)$
 (Collection-pondérée(x, AJOUT(temps)(P, t0))), t0))

Autre exemple: «EXAMEN» (TLFI)

• Action de regarder minutieusement

- En parlant de choses concrètes
- Pour mieux connaître, vérifier
- Sens particuliers
- Analyse d'une substance
- En parlant d'une personne
- Pour se faire une opinion à son sujet
- Sens médical

● Action du domaine de la pensée, de la réflexion

- En parlant de choses abstraites
- Peser les éléments d'un problème
- Mettre des principes en question
- En parlant d'une personne
- Analyse de ses propres pensées
- Étude des capacités

• Examen: un point de départ événement

- *L'examen de mercredi prochain dure 3 heures*

. **Événement** --> un élément permettant qu'il ait lieu: les questions posées par l'examineur

L'examen était difficile

- Les questions --> le support matériel (l'énoncé)

• *L'examen est à la reprographie*

• **Événement** --> résultat (les réponses)

Il y a eu beaucoup de fautes à l'examen

Les réponses --> le support matériel (les copies)

• *L'examen de maths est sur mon bureau, je ne l'ai pas corrigé*

Héritage des opérateurs applicables

• Activité --> résultat

- La *peinture* l'occupe maintenant à plein temps
- Cette *peinture* est remarquable

• Activité --> événement

- La *peinture* du plafond a duré 3 jours

• Activité --> élément permettant l'activité

- On a utilisé trois pots de *peinture*

• Activité --> lieu

- Le *marché* se tient de 9h à 13h ; le *marché* est à droite en sortant de la gare

• Objet --> événement

- Insérer le *film* dans le projecteur
- Le *film* est commencé depuis 5 minutes

• Événement duratif --> ponctuel

- La *campagne* pour le référendum sur l'Europe a marqué le point de départ des dissensions au Parti Socialiste

• Événement particulier --> générique

- L'*élection présidentielle* a eu lieu le 5 mai 2007
- L'*élection présidentielle* a lieu tous les 5 ans

Analogie synchronie-diachronie(extrait de G. Stern 1931)

13.21--> Object's names

13.211--> Material for object made of it

13.212--> Receptacle for content

13.213--> Part or constituent detail for the whole and vice versa

13.214--> Symbol for thing symbolized

13.215--> Instrument for action

13.216--> Instrument for product

13.217--> Organ for capability of perception and intellection. Object for perception

13.218--> Articles of dress or equipment for a person

13.219--> Name for concomitant circumstances

13.22--> Nomina actionis

13.221--> Action for product, result or object : a very common type

13.222--> Action for instrument or means of action

13.223--> Action for agent

13.224--> Action for place of action : very common

13.23--> Names of qualities in various uses

13.24--> Names of persons for products etc.

13.2513.251--> Place-name for action or event

13.252--> Place-name for inhabitants or frequenters

13.26--> Mental states for objects or person causing it : very common

Passage *de re-de dicto*

Un mot désignant un objet à différents moments désigne-t-il le même (**de re**) ou un autre (**de dicto**) objet

- Le *fugitif* est maintenant en prison (**de re**) [Enc, 1986]
- Le *président* est maintenant un quinquagénaire;
- avant il était septuagénaire(**de dicto**)
- avant il était quadragénaire(**de re**)
- **Cas des collections**
 - Mes *élèves* sont beaucoup plus grands que l'an dernier
 - Que ceux de l'an dernier (**de dicto**)
 - Qu'ils ne l'étaient l'an dernier (**de re**)
 - Mes *enfants* sont beaucoup plus grands que l'an dernier(**de re**)

Pluriels

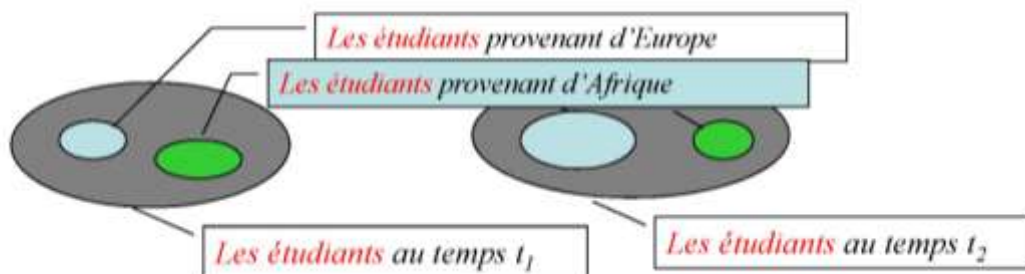
- Les *français* ont fait des progrès sur la route
- ils conduisent mieux que l'an dernier (**de re**)
- ils conduisent mieux qu'il y a 50 ans (**de dicto**)

Les français ont *hésité* pendant un siècle entre république et monarchie(**de dicto**)

- Distributif : à tout moment, chaque français hésite ?
- Collectif synchronique : chaque français a une opinion claire, mais à aucun moment il n'y a de majorité nette
- Collectif diachronique : à chaque moment il y a une majorité nette, mais elle varie fréquemment

Pluriels : Combien de collections ?

- *Les étudiants* ...
 - ... proviennent de plus en plus d'Europe et de moins en moins d'Afrique



Conclusions

- En dépit d'une multiplicité de modèles de la langue, **de nombreux problèmes centraux restent mal pris en compte**
- **Les modèles existant ne remettent pas suffisamment en cause l'arbitraire des frontières tracées**
 - Entre le linguistique et l'extra-linguistique
 - À l'intérieur de la linguistique
- **Un «bon» modèle de la langue devrait insister sur un équilibre entre contraintes de nature différente**

MAPS2 Conference
Teaching of/with Agent-Based Models in the Social Sciences
April 8th-9th, 2010 Paris, France

The Validity of Simulation Results in Social Sciences.

Christophe Sibertin-Blanc
 IRIT, sibertin@irit.fr

I. Introduction

If simulation in social sciences is “a third way for doing science” (Axelrod 1997), the practice of simulation rises questions about the nature and the conditions of the validity of the results obtained in this way. According to the scientificity principles, this validity is mainly ensured by the process carried out to obtain these results and the possibility to falsify these results or to produce incompatible results.

In collaboration with sociologists, we have formalized one of the theories of (social) organizations, the *sociology of the organized action* due to M. Crozier and E. Friedberg (Crozier 1963, Friedberg 1993), and obtained valuable simulation results in this way (Vautier et al., 2009). This paper gives an *a posteriori* account of the process that has been followed in the course of this work. This simulation-based process is sketched in Figure 3, and we believe that it is appropriate for other cases.

Any model is grounded upon a representation of the object under consideration. This representation may be implicit, and then the generality of the results provided by the study of the model is problematic, because they rely upon concepts and assumptions that are undetermined. Here, we consider cases where the underlying representation is explicit and may be considered as a theory, a well-established one in the case of the above mentioned work.

We first (section 2) present how we have formalized the sociology of the organized action – as a meta-model of the structure of organizations and a model of the rationality of social actors - and the exploitation of this formalization to produce results. This brief presentation is intended to provide an illustrative instantiation of the abstractions that are considered in the following. It also serves as a paradigmatic example of the relevance of the proposed simulation-based process, since we do not claim that it is the only way to process properly. Then (section 3), we introduce one after the other the elements that are involved in such a process and how they are related.

In conclusion, we address issues regarding the validity of simulation results. It appears that the validity of the results of a simulation-based process includes three aspects, a conceptual (or theoretical), a pragmatic and a technical one. We also claim that, in social affairs, no model may pretend to catch the entirety of a situation or a process and, in addition, the constitutive autonomy of human beings make their behavior partially unpredictable; thus, simulations can tell us, at a level of accuracy that depends on the model and its theory, what could be possible and what is presumably not, but not what will happen.

II. A Formalization of the Sociology of the Organized Action

After a brief introduction to the Sociology of the Organized Action, we give an overview of the formalization of the structure of organizations according to this theory, of means to study the analytic properties of such structures, and of the formalization of the actors' behavior that gives rise to simulations.

II.1. A brief introduction to the Sociology of the Organized Action

The Sociology of the Organized Action (SOA) (Crozier, 1964 ; Crozier *et al.*, 1977, 1995) is a well-experienced theory, acknowledged by more recent theories (Clegg et al., 1996) of organizations and simple enough to be put into a formal form. It aims at discovering and explaining the real functioning of social organizations, or so called *Concrete Systems of Action* (CSA), beyond their formal rules (in the form of organization charts, directive instructions, position descriptions, protocols and so on), and focuses on the

regularization phenomenon which ensures their relative synchronic stability. The SOA assumes that organizations are social constructs actualized by and within the interactions among its member actors. Moreover, this theory assumes that each actor behaves strategically although he has only *bounded rationality* capabilities (Simon, 1982).

The behavior of each actor is then neither totally conditioned by the organizational rules that constrain him, nor by purely individual or emotional factors. This behavior is *strategic*. Then, the induced actions aim at realizing some objectives, would it be conscious or not. To get the means to achieve his objectives, every actor seeks to have enough *power* to be able to preserve or increase his autonomy and acting capacity within the organization. This power results from the mastering of one or several *uncertainty zones* (UZ) that enable him to behave in a way that is unpredictable for other actors and consequently to set, to some extent, the *exchange rules* in his relations with others. Each uncertainty zone involves one (or several) resource for the action, and thus it is both a constraint and an opportunity. Each social actor both controls some UZs and depends on some others, so that UZs are the media of the power relationships between actors.

II.2 The structure of organizations

Purposing to draw models of social organizations, we propose a meta-model as an abstraction that catches the common concepts and properties of social organizations and can be instantiated on specific cases as models of real world organizations. The model of (the structure of) an organization is composed of instances of the meta-model's classes which correspond to the constitutive elements of the organization under consideration, and of links between these instances.

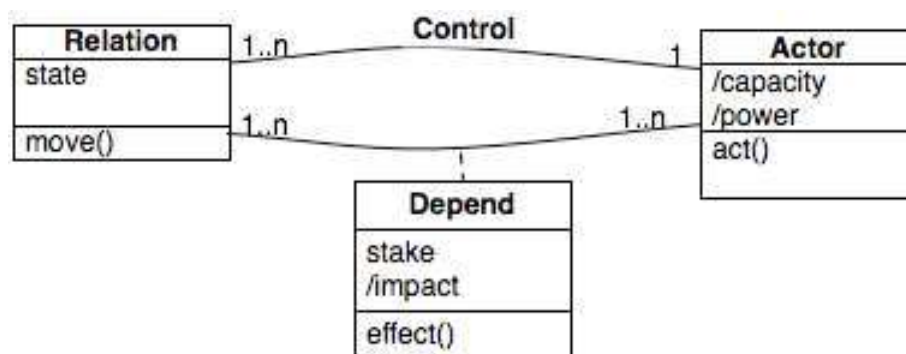


Figure 1. The meta-model of the structure of organizations

Such a meta-model is presented on figure 1, as a UML class diagram. It is composed of a class *Actor* and a class *Relation*, linked by the associations *control* and *depend*. Active entities are the *actors*, who handle the *relations*. Each relation is controlled by a unique *actor*, the only one able to move its *state*. Each actor distributes *stakes* among each relation on which he is dependent (the actor who controls a relation may also depend on this relation). Thus, he gets some *effect*, given by the *effect()* function according to the state of the relation. Consequently, an actor gets some action *capacity*, as an aggregation of the stakes and effects he receives from all the relations he depends on. He also has some *power*, as an aggregation of capacities he grants to the actors who depend on the relations he controls.

These elements will now be described in more details (see also Sibertin-Blanc et al., 2006).

Resources and Relations

Every relation is founded on an organization's resource, or a set of resources in conjunction with one another, and it is controlled by a single actor (see fig. 2). *Resources* are elements required to achieve some actions, so that their availability is necessary for some actors. A relation refers to a certain type of exchange among actors involving a specific use of the resources it is based on. Thus, the controller of a relation is in a

position to define how the resources are available to the others, and thereby, he controls to what extent they will be able to reach their goals.

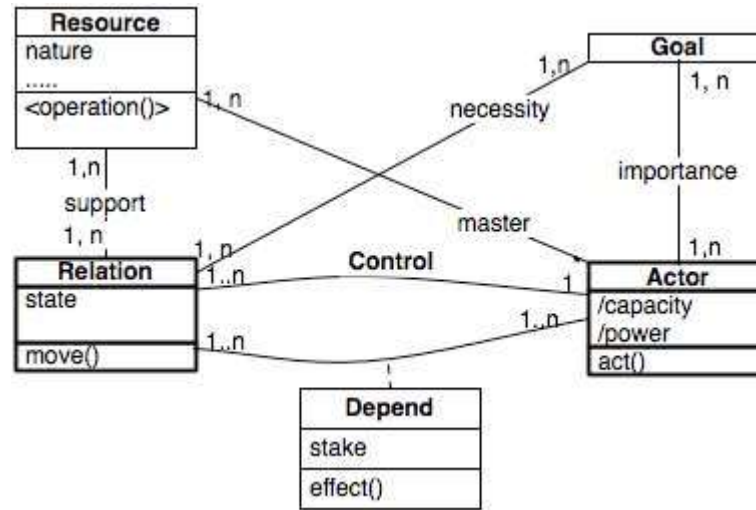


Figure 2. Resources and goals, behind the relations and the actors' stakes

The state attribute of a relation represents how the supporting resources are managed by its controller, that is, the behavior of the controller actor. It is defined inside a *space of choice* SC , representing the set of behaviors that the actor can choose to manage the relation. This space of choice is represented for any relation as the interval $SC=[-1;1]$, where the bounds -1 and 1 stand for technical or feasibility limitations, according to the nature of the underlying resources. This interval is oriented: -1 represents the least cooperative behavior of the controller, 1 represents the most cooperative behavior, and the zero value stands for a neutral behavior.

The actor who controls a relation may change his behavior by using the function *move()* that modifies the state value of the relation. An actor modifies his behavior towards a greater or lower cooperation, according to his previous behavior.

The state of a relation determines the availability of the underlying resources, i.e., how the dependent actors are granted to use them according to their goals. So, the state of a relation produces an effect on every depending actor. This effect reflects the relation's state outcome upon the capacity of the dependent actors to use the resources they need. Therefore, the greater the effect for a given actor is, the more useable the resource, and the larger his capacity of action to reach his goals. Effects take values on an arbitrary¹ scale from -10 to 10 :

worst access = -10 , ..., *neutral* = 0 , ..., *optimal* = 10 .

Effect values are given by an *effect()* function, which is defined according to the nature of the relation. The *effect_r()* function of a relation r is defined as:

$effect_r : A \times SC_r \rightarrow [-10;10]$ where A stands for the set of actors, SC_r for the space of choice of the relation r , and $[-10, 10]$ is the range of action capacity.
 $(a_1s_r) \mapsto effect_r(a_1s_r)$

Actors and their Stakes

Actors and resources are defined in relation one to another, in a dialogical way: something is a resource if and only if some actors depend on it, and someone is an actor if and only if he masters some resources, and thus some relation. The SOA assumes that actors are strategic: they have goals that lead them to find some ways to reach them. According to his goals and the necessity of related resources to reach these

¹ The scale we chose is obviously arbitrary ; since results of analysis only make sense in comparison one to another, or some simulation results, it is no longer an issue.

goals (see fig. 2), each actor distributes *stakes* on the relations. The more valuable the resource is for an actor, the higher his stake on the supported relation.

Stakes are represented by numerical coefficient, on an arbitrary scale from 0 to 10:
null = 0, *negligible* = 1, ..., *significant* = 5, ..., *critical* = 10.

Action Capacity and Power of the Actors

Defining the *state of an organization* as the vector of every relation state, each state of an organization allows to consider the aggregation of the effects of the relations on an actor, over all the relations he depends on, weighted by the stakes he puts on these relations. Such a quantity depicts the overall ability for an actor to gain access to the resources he needs to reach his goals, weighted by the relative importance of these resources regarding his goals. It measures, for an actor, his *action capacity*.

Under the hypothesis that there are no interferences between resources uses², when the organization is in the state $s = (s_{r1}, \dots, s_{rn})$, this action capacity is defined as:

$$\text{action capacity}(a,s) = \sum_{r \in R} \text{stake}(a,r) \times \text{effect}(a,s_r)$$

By controlling some relations, any actor contributes to the action capacity of the actors who depend on these relations. The global influence of an actor on the action capacity of the others, i.e. how much action capacity he gives to others, fits the concept of *power*, a core concept in the SOA theory. The power exerted by an actor a , in a state s of the organization, is defined as:

$$\text{Power}(a,s) = \sum_{r \in R, a \text{ controls } r} \sum_{b \in A} \text{stake}(a,r) \times \text{effect}(b,s_r)$$

This meta-model of organizations is completed by a couple of elements (e. g. constraints between the states of relations and solidarities between actors), but it is not necessary to elaborate on them.

II.3 Static analysis of an organization

The underlying mathematical structure of an organization allows us to define and compute the value of many *indicators* that are meaningful for the sociologist. Some of these indicators are structural - as they are not related to a specific state of the organization - and they bear either on a relation, on an actor (for instance, the relevance or the strength of a relation, the autonomy or the potential power of an actor), on relationships between an actor and a relation or on the whole organization (for instance the sum of the above indicators). Other indicators are state-dependent and also related to various kinds of elements.

The actors of an organization are related in many ways, especially by their links of solidarity and the links of dependency. For each type of such links, one can give a view of an organization as a social network, where the node actors are connected by these links and specific indicators may be used to label the nodes or the edges. These network could be analyzed using the numerous techniques and tools developed for social networks (Jackson, 2008).

The action capacity function defines, for each state of an organization, a *social configuration* which is the vector of the actors' satisfactions (and the same holds for the power and for other concepts such as the satisfaction and the influence). Exploring the possible configurations of an organization is very instructive as it allows to discover all its potential regularizations. Some of these configurations are particularly interesting because they fulfill some property such as to maximize or minimize the satisfaction of a given actor or the *global satisfaction* (i.e. the sum of all the actors' satisfactions), to be Pareto optima or Nash equilibria and so on. This exploration enables for example to know whether it is difficult for the actors to cooperate, or to detect structural conflicts between two actors when any configuration which provides one of the actors with a high satisfaction also provides the other with a low satisfaction and conversely. The *SocLab* environment computes many such interesting configurations and allows to interactively search all of

² When there are interdependencies among the impacts of the relations on an actor, they must be aggregated in a specific way; this is a matter of preference aggregation (Grabisch *et al*, 2003) well studied in the decision theory.

them; each configuration is detailed by giving the states of the relations and the amount (in value and in proportion) of capacity and satisfaction given by each actor to others by means of the controlled relations or of the solidarities.

II.4 The behavior of social actors

According to the SOA, organizations are social constructs and the social actors produce the rules that make the regularization to take place. The SOA assumes that any actor has a *strategic* behavior: by means of the relations that he controls, he seeks to put the organization in such a state that he can reach his goals, and therefore, as a *meta-goal*, to obtain enough action capacity to concretely reach them. As a consequence, that behavior is quite *stable*, as if he abided rules that, in fact, he produces. This strategic behavior is framed by a *bounded rationality* (Simon 1982) due to cognitive limitations and the opacity of social affairs. Finally, this behavior is most often *cooperative*: cooperation is required for the organization to work in a proper way, and every actor recognizes his interest in this proper working.

Thus, an organization can be seen as a *social game*, where every actor exerts his power (*i.e* set the state of the relations he controls) with the purpose to put the game in a state where he obtains a satisfactory action capacity. At each step of the game, every actor chooses an action that consists in changing the states of the relations he controls, and these changes modify the satisfaction obtained by each actor. In social games, in opposition to games considered in economics, the purpose is not to maximize the actors' action capacities, but to reach a steady state in which every actor has a satisfactory action capacity, and does not modify the relations' states anymore. Reaching a stationary state means that the actors adapted their behaviors one to another in a stable equilibrium, a regularized functioning of the organization.

To compute how an organization could be regularized, we can conduct social game simulations by providing the actors with a rationality that is compliant with the SOA assumptions and likely from the cognitive and social standpoints. The *SocLab* environment implements such a model of the actors' rationality as a process of trial and error based on a self-learning rules system (Sutton, 1998). When the simulation reaches a stationary state, we consider that a possible regularized state of the organization is computed, as an emergence of the interactions between the actors' learning processes. A sensitivity analysis module allows to study the influence of the simulation parameters upon the simulation results (El Gemayel & *al.*, 2009), and the development of facilities for the statistical data analysis (PCA, multiple correlations, ...) of the results of simulations and sensitivity analyses is in progress.

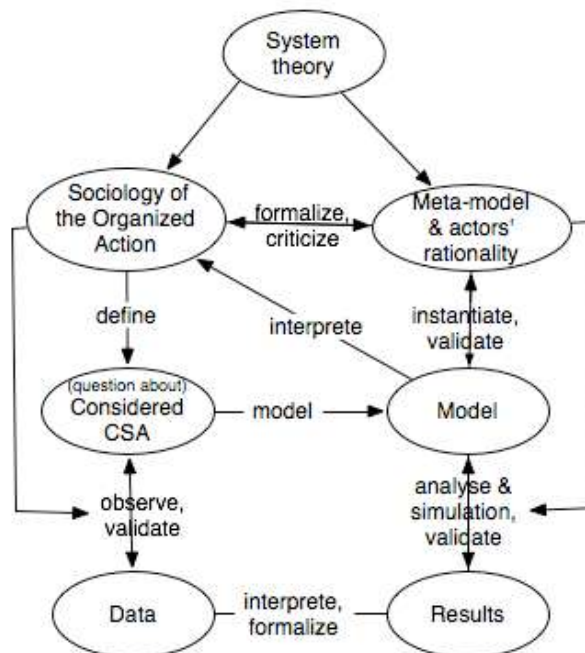


Figure 3. The simulation-based process

III. The formalization-based process

We will now comment the steps of the process shown in figure 3.

III.1. The discursive theory and its objects

We consider a theory as a conceptual framework allowing to produce specific discourses about some objects of the world.

These "objects of the world" are the things the theory is talking about and they constitute the *application domain* of the theory. Any theory has to delineate the field of its relevance and, in our case, the Concrete Systems of Action are the application domain of the SOA. The discourses of a theory about these objects is specific in as much as it aims at providing answers to specific questions about these objects; a theory does not claim to say every thing it could be possible to say about an object. For instance, the SAO intends to explain why the people of a CSA behave as they do.

To support the discourses that it allows to produce, a theory provides definitions of abstract or concrete concepts, including attributes and operations associated to these concepts, and how they are related the ones to others. An essential property of these discourses is to be expressed in natural languages, suffering from their ambiguity but also benefiting from their huge expressivity to carry a large amount of sense. In the content of a discourse of the theory about an object, we will need to distinguish the *definition part* that identifies the object under consideration, delineates its frontier and characterizes its constitutive elements, from the *analysis part* that includes the knowledge that the theory provides about this object.

In Figure 3, the "considered object" item (a CSA) corresponds to a merge of an object of the world that is considered and a definition discourse of the theory about this object, considering that the objects of the theory's application domain are established as such by the discourses about them. But our simulation-based process is also compliant with a more realistic approach that would keep a clear distinction between the objects and the discourses about them: in Figure 3, the object could be in the background, between the discourse and the model.

III.2. The theory of systems and the formalization

To formalize a theory comes down to define a formal language - that is a language including no ambiguity and allowing to make well-controlled inferences and computations - and to give means to translate the natural language discourses of the theory into this formal language. Then, a *model* is a translation of a discourse of the theory about an object of the world that is expressed into the formal language, while an *interpretation* is a converse translation of a model (expressed by means of the formal language) into a discourse of the theory.

A translation from one language into another one supposes the existence of a referential universe that includes the referents that are denoted by the terms of both the natural language discourses and the formal models. In our case, the adopted referential universe is the theory of systems, since both a CSA and its model may be viewed as systems.

From the system theory, we retain the idea that a system includes a structural aspect (its constitutive elements and their relationships) and a behavioral aspect (what are the possible or necessary changes in the state or the structure of the system), and a formalization language must provide means to express the two.

The formalism to describe the structural aspect can be expressed as a meta-model that includes the three types of entities that are constitutive of any system together with relations between them: entities (or classes) corresponding to *resources*, the passives entities that are processed by the working of the system; classes corresponding to processes (or *actors*), the active entities that have some energy at their disposal and can perform some work; and also classes corresponding to the *transitions* (or operation) that are the units of coherent changes of the organization. In our case, the meta-model of CSAs (cf. figure 1) includes a single type of resources - the relations -, a single type of processes - the social actors -, and it is useless to show the transitions - the move of the state of a relation by its controller actor.

As for the formalization of the behavioral aspect, it is defined by algorithms that determine in which cases one (or several) actor may perform some transitions involving resources and the effect of these performances. The simulation algorithm briefly presented in section II.4, is a parameterized model of the

actors' rationality to determine their behaviors in a CSA. The definition of the formal language may cause some criticisms of the discursive theory. It could be because the formal language can not deal with some inconsistencies of the theory. It is most likely because some points are implicit in the definition discourses of the theory while the formal language needs that they are taken into account, e. g. to provide a cause (or an explanation) for an effect mentioned by the theory. (For example, in the Schelling theory of segregation (Schelling, 1978), it is implicit that any individual can move toward any free position, while in any formal implementation, individuals have to know which are the free positions where they can move. This latter kind of criticisms are in fact an enrichment of the theory.

The theory may also be enriched by the definition of properties that characterize a model or some of its states as the ones presented in section II.3 (see III.5 below about the results), to the extent that they can be interpreted in the terms of the theory; the quality of these enrichments is a good measure of the relevance of a proposed formalization.

III.3. The model of an object

A model of an object that belongs to the domain of the theory is a translation of a discourse of the theory about this object that is expressed in the terms of the formal language. Thus, a model is a symbolic entity that includes (structural part) an instance of the meta-model that describes the structure of the object and (behavioral part) algorithms that describe the behaviors of its actors. There are also models that do not correspond to any observable object and are just formal construct that conform to the formalization language.

Each drawing of a model of an object can be considered as an experience that validates (or not) some qualities of the formalization.

For example, a good property of a formalization is that there is a single way to translate a discourse of the theory into a model of the formal language (orthogonality of the concepts); on the other hand, a model usually gives rise to several interpretations, either because the natural language allows semantically equivalent discourses or because any formalization entails some simplification or abstraction. As a consequence, it is expected that the model of an interpretation is the same as the initial model, while a discourse is only one of the interpretations of its model.

The easiness to produce a model that accurately catches the meaning of a discourse is another kind of validation of the formalization of the theory. It can be the case that the elaboration of the model of an object requires to acquire precise data about some elements of the object, while the theory put up with approximate values. The practical possibility to obtain these values is a validation of the applicability of the formalization.

III.4. The data

A theory goes with a methodology and tools to investigate the objects of its domain and acquire some experimental knowledges about them, knowledges which are usually called "data". The easiness to apply this methodology and to use these tools, together with the quality (consistency, accuracy, relevance, ...) of this knowledge is a validation of the discourse about the object and, in an indirect way, of the theory itself.

III.5. The results

A formalization also goes with methodologies and tools to investigate the properties of models.

Thanks to the mathematical structure of its models, a formalization allows to define techniques for their analytical study, as the ones presented in section II.3. Once the structural part of a model is implemented as a software entity, the implementation of these techniques allows to explore properties of the model. Some of these techniques propose a formal definition for concepts of the discursive theory, with the possibility to associate a quantitative or qualitative value to these concepts, e.g. the relevance of a relation in a CSA. Other techniques propose to define new concepts that do not belong to the theory but may be considered as extensions, e.g. the cooperativity of a CSA, that is the facility for the actors to find how they could cooperate for their mutual benefice and the one of the whole organization.

The behavioral part of a formalization is usually implemented as a generic simulator of models that produces simulation results, i. e. the observations of the value of variables in the course and at the end of

simulations. These simulation results are most often quite easy to produce and very numerous, and they need some processing to be presented in an easy-to-understand shape (visual display) or to highlight relevant properties (sensitivity analysis, statistic data processing or data mining).

A distinctive property of results is that they are more precise than the data; the values belong to a (qualitative or quantitative) ordinal scale and thus are easy to compare. These comparisons may be done by considering either the raw values or, especially when the scale of value is arbitrary, the relative values, that is the position of a value within its range $(\text{value} - \text{minimum_value}) / (\text{maximum_value} - \text{minimum_value})$.

IV. The validity of a formalization and of its results

The right side of the simulation-based process presented in the previous section involves 3 kinds of epistemic activities:

- conceptual (or theoretical) activities for the building of a formal language that catches the main concepts of a theory and support the translations between models and discourses, and for defining tools for the study of models (analytical techniques and simulation data to be observed);
- pragmatic (or experimental) activities, which involve methodological issues, for elaborating the model of an object and for producing meaningful results by the exploration of the model;
- technical activities for the realization of a computer application that, for any model of the formal language or for a specific model, enables its users to implement the model as a software object, to launch simulation, to collect analytical and simulation data, and to process these data.

The performance of each of these activities may be checked against specific rules about their good execution. These activities may be viewed as experiments that produce the formalization, models and results of a simulation-based process; these deliverables can be put to the test, and thus validated (or not), in various ways.

Concerning the results, the quality of the interpretable knowledge that they bring. Concerning the models, their elegance (a property which is very important in modeling matters) and their catching of the essential elements of the definition discourse of the theory about the object under consideration.

Concerning the formalization, the main test is the enabling to produce a number of successful models that also produce successful results. Another test is the knowledge that brings the techniques for the analytical study of models. Other criteria for the assessment of a formalization concern its structure (scarcity and orthogonality of its concepts, expressivity, etc.) or its practicability (the possibility to acquire the data to produce models).

Concerning the worth of the knowledge brought by modeling and simulation, we have to distinguish the objects that can be observed from the objects that are fictive.

When the model concerns an observable object, the results usually provide more precise values than the data with which they are compared. Whatever the scales of value, precise quantifications allow on the one hand to compare the rates of various features of a model and, on the other hand, to compare models (of the same object or of different ones) by comparing the rates of these features. These comparisons may be applied to the simulation results as well as to the properties defined by the analytical techniques. Models also have an explanatory power concerning the key elements of the structure of the object and its internal processes that produce the observed data.

A model and its simulations concern a fictive object if the model's structure corresponds to none observable object of the world or if the parameters of the simulations (especially the initial state) correspond to none observable state of the object. In this case, the question is whether the results of the model are credible predictions (Troitzsch, 2009). Of course, the credibility of the results depends on the extent of the validation of the formalization. However, we have to remind that a model must not be mistaken for the modeled object. A model of an object is always a simplified view of this object, that is intended to improve its understandability and to bring valuable knowledge. Thus, the proper use of results necessitates their interpretations in the field of the social reality. Moreover, considering social affairs, we have to deal with the potentially unforeseeable behavior of human beings. Thus, in this domain, predictions are only plausibilities.

V. Conclusion

This paper discusses the epistemological nature of the relationships neither between models and discourses nor between discourses and objects of the world, while it is an crucial issue about nature of the results from a scientific point of view; see (Phan et al., 2009) for this matter.

The paper considers the case of the formalization of an established theory, while noting that the theory can be enhanced by this formalization. But the simulation-based process applies in the same way when the theory is not firm and the formalization serves to test hypothesizes to build a theory. The important thing is only to consider that models are formal construct and results values in abstract ranges, and that both need interpretation to provide meaning.

References

- AXELROD, R. (1997) Advancing the Art of Simulation in the Social Sciences. In: Conte, R., Hegselmann, R., Terna, P. (eds.), *Simulating Social Phenomena*, Berlin, Springer-Verlag, pp.21-40.
- Clegg S.R., Hardy C, Nord W R (1996). *Handbook of organization studies*, Sage, London.
- Crozier M (1964). *The Bureaucratic Phenomenon*,. University of Chicago Press, Chicago.
- Crozier M., Friedberg E (1977). *L'acteur et le système: Les contraintes de l'action collective*, Paris, Seuil.
- El-Gemayel J., Sibertin-Blanc C., Chapron P. (2009). Impact of tenacity upon the Behaviors of Social Actors. In Alexis Drogoul, Benoit Gaudou, Nicolas Marilleau (Eds.), *Applied Agent based simulator Engineering for Complex System study*, Nagoya, Japan, 13/12/2009.
- Jackson M. O. (2008) *Social and Economic Networks*, Princeton University Press.
- Phan Denis, Varenne Franck (2010), Agent-Based Models and Simulations in Economics and Social Sciences: from conceptual exploration to distinct ways of experimenting; *Journal of Artificial Societies and Social Simulation* 13, 1.
- SCHELLING, T.S. (1978) *Micromotives and Macrobehaviour*, N.Y, Norton and Co.
- Sibertin-Blanc C, Amblard F, Mailliard M, (2006). A coordination framework based on the Sociology of Organized Action. In O. Boissier, J. Padget, V. Dignum, G. Lindemann (Eds.), *Coordination, Organizations, Institutions and Norms in Multi-Agent Systems* (pp. 3-17), Springer, LNCS 3913.
- Simon H. A. (1982). Models of bounded rationality: Behavioral economics and business organization (Vol. 1 et 2), The MIT Press.
- Sutton R.S., Barto A. G. (1998). *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA.
- Troitzsch, K. (2009), Not All Explanations Predict Satisfactorily, and Not All Good Predictions Explain, *Journal of Artificial Societies and Social Simulation vol. 12, no. 1 10*
- Vautier C., Roggéro P., Adreit F., Sibertin-Blanc C., (2009). Evaluation by simulation of the social acceptability of agricultural policies for water quality , In *International ICSC symposium on Information Technologies on Environmental Engineering*, Thessaloniki, Greece, mai 2009, Springer Verlag, p. 478-490.

Modèles formels pour reproduction des simulations à base d'agents

Christophe Sibertin-Blanc

Université de Toulouse / IRIT

2 rue du doyen G. Marty, F31042 - Toulouse cedex 9

sibertin@univ-tlse1.fr

Nouvelles perspectives en Sciences Sociale 5, 2, p.135 - 149, 2010.

Résumé : Du fait de la dimension expérimentale des résultats obtenus par simulation informatique, la validation de ces résultats nécessite qu'ils puissent être reproduits. Cela suppose que le dispositif concret à partir desquels ils sont obtenus, à savoir le logiciel dont les exécutions fournissent les sorties qui sont observées, puisse donner lieu à de nouvelles implantations. Bien qu'elle soit indispensable à maints égards, il s'avère que cette "réplication" des modèles de simulation est le plus souvent problématique ; cet article propose de présenter ces modèles sous le forme de systèmes afin d'en faciliter de nouvelles implantations.

Mots clés : simulation sociale, agents, expérimentation, reproduction des résultats, système, modèle formel, théorie.

Abstract. Since the results obtained by means of computer-based simulations have an experimental character, their validation requests their reproduction. These results are obtained by observing the outputs of the runs of a software entity, which is the concrete device of the experimentation; thus, it is this software simulation model that must support a new implementation. Although the replication of a model brings many advantages, in many cases it is problematic. This paper proposes to present simulation models as systems, to ease their replication.

Key words. social simulation, agents, experimentation, replication, system, formal model, theory.

I. Introduction

Dans le domaine des sciences sociales, la simulation à base d'agents est une pratique qui se développe assez largement et est reconnue comme "a third way for doing science"³. Elle comporte un aspect déductif - le modèle qui donne lieu à des simulations est un objet formel sur lequel des inférences et des calculs logiquement contrôlés peuvent être réalisés. Elle comporte aussi un aspect expérimental puisque les simulations sont destinées à acquérir des résultats nouveaux sur le comportement de ce qui est modélisé ; ces résultats ne peuvent pas être obtenus de façon analytique, du fait des phénomènes d'émergence qui résultent de la multiplicité des (types d') interactions entre les éléments qui sont en jeu⁴. Se pose donc le problème de la reproduction des résultats de ces simulations, qui ne peuvent être considérés comme valides que s'ils sont reproduits : "*unreplicated simulation models and their results can not be trusted - as with other kinds of experiment, simulations need to be independently replicated. [...] aligning models is very difficult, but very revealing*"⁵.

Mais de quelle nature est cette réplication, ou reproduction ? Tout processus visant à obtenir des résultats scientifiques au moyen de simulations met en jeu quatre éléments. Il y a d'abord, dans un certain domaine, l'objet qui a été repéré comme digne d'intérêt et constitue la *cible du modèle*. Le *modèle conceptuel* de cette

³ Robert Axelrod, "Advancing the Art of Simulation in the Social Sciences", dans Rosaria Conte, R. Hegselmann et P. Terna (dir.), *Simulating Social Phenomena*, Berlin, Springer-Verlag, p. 21-40, 1997.

⁴ Dessalles, Jean-Louis, Jean-Pierre Müller et Denis Phan, "Emergence in Multi-agent Systems : Conceptual and Methodological Issues", dans Denis Phan et Frédéric Amblard, (dir), *Agentbased Modelling and Simulation in the Social and Human Sciences*, Oxford, The Bardwell Press, p. 327-356, 2007.

⁵ Edmonds, Bruce et David Hales, "Replication, Replication and Replication: Some Hard Lessons from Model Alignment", *Journal of Artificial Societies and Social Simulation*, vol. 6, no. 4, 2003.

cible en est la représentation sur laquelle se fonde son étude, représentation plus ou moins bien fondée sur une théorie des sciences sociales et qui est, en tout état de cause, exprimée en langage naturel. Le *modèle formel* est une traduction du modèle conceptuel qui a pour fonction d'en évacuer les ambiguïtés (en faisant certains choix réducteurs) et d'être une description, ou spécification, complète et consistante du dispositif dont l'exécution sur ordinateur produira des résultats de simulation. Vient enfin le *modèle implanté* qui sera le support opérationnel des simulations réalisées sur un ordinateur⁶. Le passage du modèle conceptuel vers le modèle formel et de ce dernier vers le modèle implanté résulte d'une forme de traduction, chacun d'eux étant formulé dans un langage qui lui est propre, avec les procédés de construction de sa syntaxe, le pouvoir d'expression de sa sémantique et ses tests de cohérence interne. En sens inverse, c'est un processus d'induction / généralisation qui permet de formuler les résultats relatifs au modèle conceptuel à partir de ceux relatifs au modèle formel, et ces derniers à partir d'observations des sorties de simulations du modèle implanté. C'est donc le modèle implanté qui détermine les conditions pratiques de l'expérience qui a permis d'obtenir des résultats, et renouveler l'expérience nécessite de répliquer le modèle implanté, de réaliser une nouvelle implantation du modèle formel.

De nombreux auteurs ont signalé combien cette réplique est difficile. D'une part, les techniques de vérification du génie logiciel qui permettent de s'assurer qu'une implantation est correcte, c'est à dire qu'elle respecte ses spécifications, ne sont pas applicables ici : l'essence d'une simulation est de produire des résultats nouveaux, et on ne dispose donc pas de l'oracle qui permettrait, en fonction des entrées, de comparer les sorties obtenues à celles qui sont attendues et devraient être produites⁷. D'autre part, si les articles publiés fournissent généralement une bonne présentation du modèle conceptuel, ce qui est indispensable pour que le lecteur se fasse une idée assez claire de ce dont il s'agit, le modèle formel, qui hérite nécessairement d'une partie de la complexité de la cible du modèle, est plus difficile à présenter si bien que l'on ne dispose pas de suffisamment d'informations pour en dériver une nouvelle implantation qui produise les mêmes résultats que ceux annoncés.

On pourra se référer par exemple à Wilensky⁸ pour une synthèse des bénéfices de la réplique d'un modèle, sur les difficultés rencontrées et sur les questions de validation, à savoir les critères permettant de dire qu'une réplique d'un modèle formel confirme ou non les résultats antérieurs associés à ce modèle. Cet article n'abordera pas davantage la question centrale de la "validité" des résultats et de leur statut épistémologique, qui met en cause les relations entre la cible du modèle et le modèle conceptuel et celles entre le modèle conceptuel et le modèle formel (qui en est toujours une simplification), ni la légitimité du processus de production des résultats portant sur le modèle formel à partir des sorties du modèle implanté. Nous nous contenterons de proposer une façon de structurer la présentation de modèles formels, en d'autres termes un formalisme (qui est composite, du fait de la diversité des dimensions à prendre en compte), qui vise à favoriser la réplique de leur implantation. C'est la complétude de la description d'un modèle formel qui pose problème, et non pas sa consistance qui, de par la nature du modèle, peut être soumise à des tests qui vérifient l'absence de contradictions.

Introduire une façon de structurer un objet nécessite de passer à un niveau d'abstraction supérieur pour catégoriser les éléments entrant dans sa composition. Ici, nous nous baserons sur des concepts très généraux de la théorie des systèmes. Ce choix est justifié dans la mesure où la cible du modèle peut être analysée

⁶ Drogoul, Alexi, Diane Vanbergue et Meurisse Thomas, "Multi-agent Based Simulation: Where Are the Agents?", dans Jaime Sichman, François Bousquet et P. Davidsson (dir) *Proceedings of MABS 2002 Multi-Agent-Based Simulation*, Lecture Notes in Computer Science 2581, p. 1-15, Springer-Verlag, 2003.

⁷ Amblard, Frédéric, Juliette Rouchier et Pierre Bommel «Evaluation et validation de modèles multi-agents», dans Amblard Frédéric, Phan Denis (dir.) *Modélisation et simulation multi-agents, applications pour les Sciences de l'Homme et de la Société*, Londres, Hermes-Sciences & Lavoisier, p. 103-120, 2006

⁸ Wilensky, Uri et William Rand, "Making Models Match: Replicating an Agent-Based Model", *Journal of Artificial Societies and Social Simulation* vol. 10, no. 4, 2007.

comme un système, le propre d'un modèle de simulation à base d'agents étant précisément de se calquer sur la structure de cette cible. Si un modèle formel est donc la description d'un système, il devra comporter une description d'une part de sa *structure*, ses éléments constitutifs et les relations qu'ils entretiennent, et d'autre part de son *comportement*, ou fonctionnement, à savoir comment il peut ou doit évoluer.

II. La structure d'un modèle formel et son espace d'états

Pour modéliser la structure d'un système, il faut donc identifier, à partir du modèle conceptuel, les éléments des trois catégories d'éléments constitutifs de tout système, à savoir : les entités passives, c'est-à-dire les choses de toute nature qui sont manipulées au cours du fonctionnement du système, auxquelles certains traitements sont appliqués, et que l'on appellera les *ressources* du système ; les entités actives, ou *acteurs*, *processeurs*, qui disposent d'une source d'énergie et sont capables de réaliser un certain travail de façon plus ou moins autonome⁹; et enfin les *transitions*, unités de travail réalisées par des acteurs, mettant en jeu des ressources et dont l'occurrence fait évoluer l'état ou la structure du système. De façon standard dans un système à base d'agents, les entités passives et actives seront des agents décrits par la donnée de leur nom, des *attributs* qui leur sont associés et des *services* (ou fonctions, méthodes) qu'ils sont capables de réaliser (dans le cas des acteurs) ou qui leurs sont applicables (dans le cas des ressources) ; s'agissant des transitions, elles seront décrites par leur nom et la description du travail correspondant, à savoir les services des acteurs et ressources concernés qu'il s'agit d'effectuer. La description de la structure d'un modèle nécessite aussi de décrire les *relations* que les éléments de ces trois catégories entretiennent les uns avec les autres.

Ces entités peuvent être classées selon leur type, et les types des entités (de la structure) du modèle et les types de leurs relations constituent un *méta-modèle* qui peut être représenté sous la forme d'un diagramme de classe UML tel que celui de la figure 1 pour le cas, paradigmatique dans le domaine de la simulation sociale, du modèle de ségrégation résidentielle de Schelling¹⁰. Ce diagramme de classes doit être accompagné de contraintes sémantiques sur la façon dont il peut être instancié ; UML permet d'exprimer certaines de ces contraintes par les cardinalités associés aux rôles (par exemple un individu occupe un et un seul emplacement), mais il y en a bien d'autres : un individu occupe un emplacement qui lui est accessible, la satisfaction d'un individu est la valeur retournée par sa fonction calcul Satisfaction, etc.

Les types des entités du modèle et leurs relations peuvent tout aussi bien être décrits par une structure mathématique¹¹ en terme d'ensembles, de fonctions et d'axiomes ; s'agissant du diagramme de classe de la figure 1, cette structure est la suivante :

Un ensemble pour chacun des type d'entités, soit :

- un ensemble Individu ;
- un ensemble Emplacement, tel que $|Individu| < |Emplacement|$;
- un ensemble Territoire, tel que $|Territoire| = 1$;

Une fonction pour chacun des services :

- une fonction calcul Satisfaction : $Individu \rightarrow \{content, mécontent\}$;
- une fonction libérer : $Individu \times Emplacement \rightarrow T$;
- etc.

Une relation ou une fonction pour chacune des relations :

⁹ Les êtres animés ont vocation à être des acteurs et les objets à être des ressources, mais pas nécessairement. Dans un modèle de ruissellement par exemple, les gouttes d'eau qui consomment de l'énergie potentielle pour se déplacer seront modélisées comme des acteurs, à moins qu'elles ne soient agies par un processeur "gravitation" qui exerce l'ensemble des forces d'attraction ; réciproquement, un humain sans aucune initiative et intégralement instrumenté par d'autres aura le statut de ressource.

¹⁰ Thomas S. Schelling, *Micromotives and Macrobehaviour*, N.Y, Norton and Co, 1978.

¹¹ Plus précisément une classe d'objets mathématiques.

- une fonction appartenir : Emplacement | > Territoire ;
- une fonction injective occuper : Individu | > Emplacement ;
- une relation accessible C Individu x Emplacement ;
- une relation (symétrique) voisin C Emplacement x Emplacement ;

Une fonction pour chacun des attributs des entités :

- une fonction couleur : Individu | > {blanc, noir} ;
- etc.

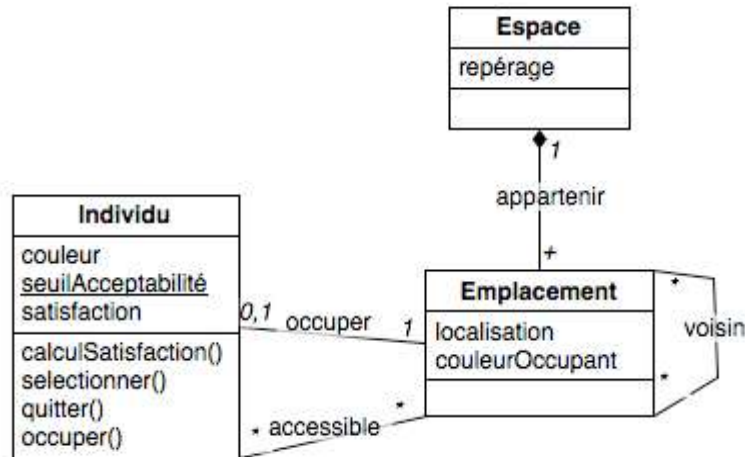


Figure 1. Diagramme du classe UML du méta-modèle de la ségrégation résidentielle selon Schelling

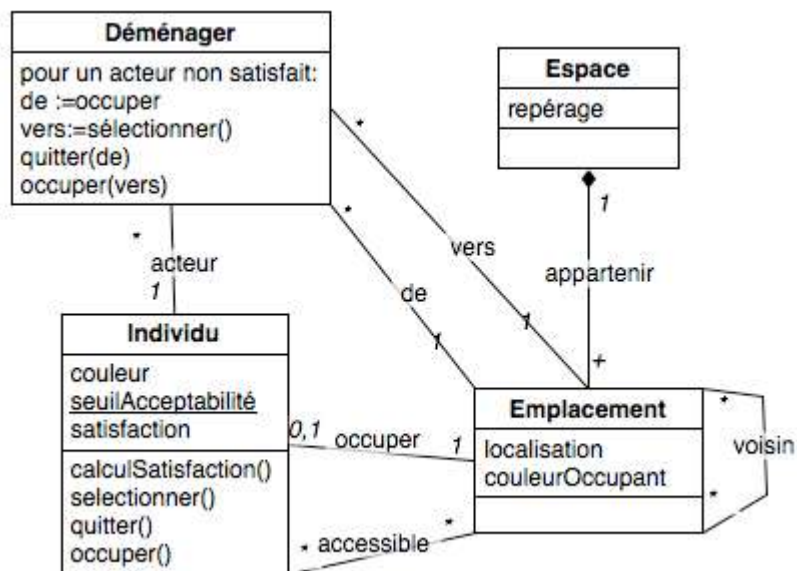


Figure 2. Méta-modèle de la ségrégation résidentielle selon Schelling faisant figurer les transitions

Concernant les transitions, leurs types peuvent soit figurer sur un diagramme de classe tel que celui de la figure 2, soit être décrits séparément.

Ce que nous appellerons *la structure du modèle* est alors la donnée de :

- l'ensemble des instances des différents types d'entités, ici l'ensemble des Territoires (réduit à un unique élément), l'ensemble des Individus et l'ensemble des Emplacements ;
- la spécification des services de ces entités, par exemple sous la forme de pré- et postconditions;
- la description de chacun des types de transitions, par exemple sous la forme d'un algorithme ou d'un ensemble de règles.

Les transitions ne sont pas de même nature que les entités, leurs instances, que l'on appelle plus communément *occurrences*, étant par nature éphémères, et même instantanées dans le cas des systèmes discrets dont l'état est non observable, ou indéterminé, pendant la durée de l'exécution d'une transition. Cependant, ce sont les entités en présence qui déterminent les types des transitions qui peuvent avoir lieu, et cela justifie de les considérer comme constitutifs de la structure d'un modèle. Ceci dit, considérer que la définition des types de transitions relève de la description du comportement du modèle (cf. ci-dessous) produit, au final, la même description d'ensemble.

Ce que nous appellerons *un état du modèle* (dont la structure est préalablement définie), c'est la donnée des liens (i.e. les instances des types de relations) entre les entités ainsi que la valeur de chacun de leurs attributs¹². Un tel état peut être représenté soit sous la forme d'un diagramme d'objet UML qui soit conforme au diagramme de classe et respecte ses contraintes sémantiques, soit en explicitant, dans la représentation sous forme mathématique, les ensembles et les fonctions correspondant aux relations et aux attributs. L'ensemble des états d'un modèle constitue donc son *espace d'états*. La séparation entre ce qui relève de la structure d'un modèle et de ce qui relève de son état peut paraître arbitraire, en tout état de cause elle met en jeu des choix de modélisation ; par exemple, il est bien connu que toute relation peut tout aussi bien être représentée par une entité. Ces choix de modélisation sont le plus souvent guidés par des critères d'esthétique, dimension essentielle de toute modélisation que nous ne discuterons pas ici, nous contentant de remarquer que ces choix correspondent le plus souvent à des représentations mathématiques qui sont isomorphes.

III. Le comportement d'un modèle.

La définition mathématique de la structure d'un modèle permet de définir des propriétés internes de ce modèle et de les étudier *analytiquement*. Par exemple, dans le cas du modèle de ségrégation de Schelling, on peut calculer la proportion d'emplacements inoccupés (propriété purement structurelle, indépendante de tout état), ou bien le taux de ségrégation ou la taille des secteurs homogènes pour un état donné.

La réalisation de simulations nécessite que soit définie la dynamique du système, c'est-à-dire comment son état ou sa structure peuvent évoluer.

La structure d'un système permet de définir un ensemble *d'actions*, à savoir la réalisation par un (ou plusieurs) acteur d'une transition mettant en jeu certaines ressources, dont les occurrences dans un état donné du système sont des *actes*. Pour simplifier les choses, nous considérerons ici que les actes modifient l'état du modèle mais ne modifient pas sa structure ; ce n'est pas toujours le cas, voir par exemple la modélisation des configurations sociales de la sociologie de l'imprévisible¹³, mais la prise en compte d'actions qui font évoluer la structure d'un modèle ne change rien à sa présentation, elle est simplement un peu plus compliquée techniquement.

La *fonction de transition* d'un modèle, qui pour chaque état de son espace d'états indique les (ensembles d') actes qui peuvent se produire et le nouvel état qui résulte de leur réalisation respective, décrit l'ensemble des comportements envisageables du système. Depuis un état initial donné, toute séquence finie (d'ensembles)

¹² Si l'on a affaire à un système continu, non discret, l'état du système doit aussi prendre en compte l'état interne des occurrences de transitions qui sont en cours.

¹³ Grossetti, Michel et Christophe Sibertin-Blanc, Un méta-modèle des configurations de la sociologie de l'imprévisible, dans Denis Phan (dir.) *Ontologies pour la Modélisation Multiagents en Sciences Humaines et Sociales*, à paraître, 2010.

d'actes conforme à cette fonction de transition définit une trajectoire dans l'espace d'états du modèle, depuis cet état initial jusqu'à l'état terminal. La définition de cette fonction permet de faire certaines vérifications de la cohérence du modèle, par exemple par le calcul des états qui sont accessibles depuis un état initial donné. Si modélisateur a une idée des états initiaux et terminaux qui sont intéressants à considérer, il peut s'assurer que les derniers sont accessibles à partir des premiers. Mais ce n'est généralement pas à partir de cette fonction de transition que sont réalisées des simulations.

En effet, le modèle conceptuel à partir duquel le modèle formel est élaboré comporte une description des processus qui sont à l'oeuvre dans le système cible du modèle, ou bien des hypothèses sur ces processus qu'il s'agira de tester. Pour l'essentiel, les résultats de simulation indiquent les effets de ces processus, et ce sont donc eux qu'il s'agit de formaliser pour les mettre en oeuvre dans des simulations.

Par définition des trois catégories d'entités qui interviennent dans la structure d'un modèle, ces processus sont portés par les acteurs ; chaque acteur du modèle réalise un processus, et le comportement global du système résulte des interactions entre ces processus.

Il s'agit donc dans un premier temps de décrire le processus de chacun des acteurs, en fait de chacun des types d'acteurs puisque les acteurs d'un même type suivent, par définition, le même schéma comportemental. Ce processus se déroule toujours selon le cycle :

- répéter
- mettre à jour des informations dont l'acteur dispose sur l'état du système
- sélectionner un acte (éventuellement de ne rien faire)
- réaliser cet acte jusqu'à la fin et ce sont donc les étapes de ce cycle qui doivent être décrites, en indiquant comment chaque acteur met à jour ses informations et comment il sélectionne le prochain acte qu'il va réaliser¹⁴. Il faut ensuite décrire comment les processus de chacun des acteurs sont synchronisés les uns avec les autres. Les choses sont relativement simples lorsque chaque acte est réalisé à l'initiative d'un seul acteur. Dans ce cas, s'il n'y a pas une gestion spécifique du temps, soit la simulation fait agir les acteurs de façon asynchrone, comme dans le modèle de ségrégation de Schelling, selon la boucle :

- répéter
- pour chacun des acteurs sélectionnés
- réaliser un cycle

jusqu'à la fin,

soit elle les fait agir de façon synchrone, selon la boucle :

- répéter
- pour chacun des acteurs sélectionnés
- mettre à jour les informations
- sélectionner un acte

pour chacun des acteurs sélectionnés

- réaliser l'acte sélectionné

jusqu'à la fin.

La synchronisation entre les processus des acteurs est plus délicate à décrire lorsque des actes peuvent être réalisés à l'initiative conjointe de plusieurs acteurs. La synchronisation s'opère alors non plus au niveau des cycles des acteurs comme dans le cas précédent, mais au niveau de l'étape "sélectionner un acte" de ces cycles. Nous n'avons pas étudié cette question, mais il serait certainement intéressant de disposer de mécanismes simples et généraux pour exprimer ce type de synchronisation¹⁵.

¹⁴ Les acteurs correspondant à des choses, cf. note 7, sont généralement *déterministes*, alors que ceux correspondant à des êtres animés sont le plus souvent dotés d'une certaine *autonomie*, ils choisissent leur prochain acte parmi un ensemble d'alternatives.

¹⁵ On peut remarquer cependant que selon les théories sociologiques de l'action qui mettent en avant les choix rationnels des acteurs sociaux (par exemple Coleman, Crozier et Friedberg,

Il reste enfin à caractériser l'état initial des simulations, y compris la valeur des paramètres qui interviennent dans le comportement de chaque acteur, et le critère de fin des simulations ; cela ne pose généralement pas de difficulté. Les simulations se terminent le plus souvent par la satisfaction d'un critère de convergence, l'atteinte d'un état stationnaire ou d'un état particulier, ou encore la durée de la simulation.

IV. Conclusion

On a exposé dans cet article une certaine façon de présenter des modèles formels destinés à être implantés pour donner lieu à des simulations informatiques, dans le but de faciliter la reproduction des expériences que constituent ces simulations. Cette façon de présenter constitue en fait un formalisme qui est présenté ici de façon informelle, c'est-à-dire sans définition explicite de ses niveaux lexical, syntaxique et sémantique, ce qui serait inutilement technique. Il convient toutefois de remarquer qu'un modèle formel, même complet et consistant, ne conduit pas automatiquement à des implantations équivalentes, au sens où les mêmes entrées produisent toujours les mêmes sorties. Ce serait compter sans les *biais d'implantation* qui tiennent aux choix d'implantation du programmeur (inévitables tant que le code du modèle implanté n'est pas généré automatiquement à partir d'un modèle formel dont la définition est intégralement formalisée), au langage de programmation, aux bibliothèques utilisés, au système d'exploitation de l'ordinateur ou aux caractéristiques physiques de ce dernier. A titre d'exemple, sont en cause la discrétisation d'un phénomène continu, la réalisation de calculs en virgule flottante ou selon l'arithmétique entière, le générateur de nombres aléatoires, ou encore la gestion des fils d'exécutions (les *threads*). Winsberg¹⁶ étudie certains de ces inévitables biais ; Galán¹⁷ indique comment les mettre en évidence par la réplication, ce qu'elle seule permet de faire, et de préférence en double aveugle. On peut se demander avec D. Phan¹⁸ s'il est envisageable de produire des modèles formels qui éliminent tous les biais d'implantation, c'est-à-dire dont toutes les implantations correctes soient équivalentes. Nous ne le pensons pas, de même qu'il y a le plus souvent plusieurs façons de formaliser un modèle conceptuel en un modèle formel, il peut y avoir plusieurs façons non équivalentes d'implanter un modèle formel. Le modèle formel est un objet purement symbolique (plus précisément un système de symboles), ce qui n'est le cas ni du modèle implanté ni de son environnement d'exécution. Tout au plus, un modèle formel pourrait-il être accompagné de contraintes d'implantation, faisant alors référence à un état donné de la technologie.

Dans les démarches scientifiques basées sur la simulation, il est fréquent que l'on ne s'intéresse pas à un unique système cible, mais que les résultats soient annoncés comme ayant une portée plus générale et applicables à des systèmes ou phénomènes sociaux différents ; il est évident, par exemple, que les résultats associés au modèle de ségrégation de Schelling sont indépendants de la structure concrète du territoire, par exemple sa taille. Dans ces situations, l'auteur du modèle dispose, implicitement ou non, d'une théorie qu'il applique (ou pourrait appliquer) à divers systèmes concrets pour chacun desquels il est en mesure de produire un modèle conceptuel. Et c'est la parenté entre ces modèles qui fonde la généralisation des résultats à l'ensemble de ces systèmes. Pour cerner cette parenté, et donc la légitimité de cette généralisation, on peut regarder ce qu'il en est des modèles formels de ces différents systèmes. Ce qui justifie la généralisation des résultats, c'est que ces modèles formels partagent le même métamodèle- leurs structures étant composées

Boudon, Touraine, Bourdieu pour partie), les interactions entre les acteurs sont médiatisées par des ressources, notamment celles cognitives, dont ils partagent l'accès. Dans ce cas, la décision de réaliser conjointement une action est elle-même une action, de niveau méta, que chacun réalise pour sa part, et on est ramené au cas où seuls les cycles des acteurs sont alors à synchroniser, les acteurs se coordonnant par l'intermédiaire des ressources.

¹⁶ Eric Winsberg, Sanctioning models: the epistemology of simulation, *Science in Context*, 12, p. 275-292, 1999.

¹⁷ Galán, José Manuel, Luis R. Izquierdo, Segismundo S. Izquierdo, José Ignacio Santos, Ricardo del Olmo, Adolfo López-Paredes et Bruce Edmonds, "Errors and Artefacts in Agent-Based Modelling", *Journal of Artificial Societies and Social Simulation* vol. 12, no. 1, 2009,

¹⁸ Phan, Denis, communication privée.

des mêmes types d'acteurs, de ressources et de transitions - et qu'ils ont la même dynamique, les résultats généraux portant sur la nature et l'effet des interactions que peuvent avoir des acteurs des différents types .

Pour conclure ce paragraphe, un modèle formel dont on néglige les instances de ses types d'acteurs et de ressources peut être considéré comme une formalisation de la théorie, encore une fois implicite ou non, mise en jeu pour produire le modèle conceptuel dont il est issu, et l'ensemble de ses instanciations comme le domaine d'application de cette théorie.

Remerciements. Cet article doit beaucoup à de très nombreuses discussions avec tous les participants du projet SocLab, et plus particulièrement Françoise Adreit, Pascal Roggero et Claude Vautier, ainsi qu'à des échanges très enrichissants dans le cadre du projet ANR COSMAGENS, notamment avec Denis Phan.

Bibliographie

- Robert Axelrod, "Advancing the Art of Simulation in the Social Sciences", dans Rosaria Conte, R. Hegselmann et P. Terna (dir.), *Simulating Social Phenomena*, Berlin, Springer-Verlag, p. 21-40, 1997.
- Dessalles, Jean-Louis, Jean-Pierre Müller et Denis Phan, "Emergence in Multi-agent Systems : Conceptual and Methodological Issues", dans Denis Phan et Frédéric Amblard, (dir), *Agentbased Modelling and Simulation in the Social and Human Sciences*, Oxford, The Bardwell Press, p. 327-356, 2007.
- Drogoul, Alexi, Diane Vanbergue et Thomas Meurisse, "Multi-agent Based Simulation: Where Are the Agents?", dans Jaime Sichman, François Bousquet et P. Davidsson (dir) *Proceedings of MABS 2002 Multi-Agent-Based Simulation*, Lecture Notes in Computer Science 2581, p. 1-15, Springer-Verlag, 2003.
- Edmonds, Bruce et David Hales, "Replication, Replication and Replication: Some Hard Lessons from Model Alignment", *Journal of Artificial Societies and Social Simulation*, vol. 6, no. 4, 2003.
- Galán, José Manuel, Luis R. Izquierdo, Segismundo S. Izquierdo, José Ignacio Santos, Ricardo del Olmo, Adolfo López-Paredes et Bruce Edmonds, "Errors and Artefacts in Agent-Based Modelling", *Journal of Artificial Societies and Social Simulation* vol. 12, no. 1, 2009,
- Grossetti, Michel et Christophe Sibertin-Blanc, Une formalisation de la dynamique des configurations sociales : dialogue entre un sociologue et un informaticien., dans Denis Phan (dir.) *Ontologies pour la Modélisation Multi-agents en Sciences Humaines et Sociales*, à paraître, 2010.
- Thomas S. Schelling, *Micromotives and Macrobehaviour*, N.Y, Norton and Co, 1978.
- Vautier, Claude, Pascal Roggero, Françoise Adreit, Christophe Sibertin-Blanc, "Evaluation by simulation of the social acceptability of agricultural policies for water quality", actes de *International ICSC symposium on Information Technologies on Environmental Engineering*, Thessaloniki (Grèce), mai 2009, Springer Verlag, p. 478-490, 2009.
- Wilensky, Uri et William Rand, "Making Models Match: Replicating an Agent-Based Model", *Journal of Artificial Societies and Social Simulation* vol. 10, no. 4, 2007.
- Eric Winsberg, "Sanctioning models: the epistemology of simulation", *Science in Context*, 12, p.275-292, 1999.

