Engagements, intentions et jeux de dialogue

Philippe Pasquier, Brahim Chaib-draa

Département Informatique-Génie Logiciel, Université Laval
Sainte-Foy G1K 7P4 - Canada (Québec)

http ://DAMAS.ift.ulaval.ca
courriels : pasquier@iad.ift.ulaval.ca, chaib@iad.ift.ulaval.ca

RÉSUMÉ. *L'objectif de cet article est de présenter, tout en contribuant à les compléter, les approches conventionnelles pour le dialogue entre agents, en particulier les approches par les jeux de dialogues. Pour ce faire, nous présentons un cadre interactionnel conventionnel complet (DIAGAL) pour la manipulation de la couche publique (constituée d'engagements sociaux). Nous introduisons et discutons le paradigme qui nous semble sous-jacent (et généralement implicite) aux approches conventionnelles récemment proposées pour la communication agents. Nous indiquons comment un lien pourrait être établi entre les aspects publics (représentés par les engagements sociaux) et les aspects privés aux agents (essentiellement leurs états mentaux, résultant en intentions). Ce faisant, nous dégageons un certain nombre de problèmes théoriques que la communauté SMA(Systèmes Multi-Agents) devra investir pour que les nouvelles approches conventionnelles des communications entre agents soient convaincantes sur le plan théorique et utilisables concrètement.*

ABSTRACT.

MOTS-CLÉS : *communication entre agents, jeux de dialogue, intention, engagement sociaux*

KEYWORDS:

1. Introduction

Les technologies agents et multi-agents permettent de concevoir et de développer des applications complexes. La caractéristique fondamentale de ces systèmes dans le paradigme actuel de l'informatique répartie est l'habileté des agents à communiquer entre eux de manière utile à leurs objectifs tant individuels que collectifs. Pour ce qui est des communications, la communauté SMA se concentre depuis quelques années sur l'élaboration d'un cadre interactionnel standard.

Les principaux langages de communication agent actuels, KQML [FIN 94] et FIPA-ACL [FIP 02], sont basés sur la théorie des actes de langage [SEA 69]. La sémantique des messages est formulée en terme d'états mentaux, aspects privés aux agents [FIN 98]. Le dialogue est censé émerger de l'enchaînement des productions d'actes issus des intentions de chaque agent via la reconnaissance et le raisonnement sur les intentions des autres. Cette approche, dite « mentaliste », a été critiquée [MOU 97, SIN 98]. Elle pose, entre autres, le *problème de la vérification sémantique* : c'est-à-dire que les agents doivent pouvoir vérifier que les autres agissent bien conformément à ce que les dialogues tenus indiquent[1]. Pour que la sémantique des messages soit vérifiable, il faudrait avoir accès aux états mentaux privés des agents ce qui n'est généralement pas possible. Un second problème majeur posé par cette formulation est *le problème de l'hypothèse de sincérité* nécessaire à la définition d'une telle sémantique. Cette hypothèse est jugée trop contraignante par la communauté SMA. Elle interdit notamment d'envisager correctement certains types de dialogues dans des domaines où telle hypothèse ne saurait tenir. C'est le cas ,par exemple, pour les dialogues de négociation dans le commerce électronique [DIG 00].

D'une manière générale, le comportement des agents cognitifs est basé sur les attitudes mentales privées de l'agent. Or, les agents sont habituellement définis en termes d'autonomie, d'interopérabilité et de leur capacité à atteindre ensemble des buts communs. Le point de vue mentaliste, centré sur l'agent-individu, n'aide pas à définir une sémantique pour ce qui est commun et partagé dans un groupe d'agents. Finalement, dans ce cadre mentaliste, la communication agent souffre d'un manque de sémantique formelle concise et universellement acceptée, de sorte que la communication agent est restreinte à des domaines précis dans des environnements qui ne sont pas ouverts.

Un certain nombre de solutions à ces problèmes ont été récemment proposées dans lesquelles le caractère social des agents est mis de l'avant [SIN 98, COL 98, MOU 97]. La notion d'engagement social a rapidement émergé comme permettant de capturer le niveau public du dialogue. De nombreux auteurs ont proposé des approches sociales des communications agents introduisant une couche conventionnelle exprimée en terme d'engagements sociaux [SIN 00, COL 00, FLO 01, AMG 00]. Ces approches permettent d'améliorer la vérifiabilité (voir section 3.2), de lever l'hypothèse de sin-

1. À ne pas confondre avec la vérification sémantique informatique qui consiste à vérifier que les agents implémentent la sémantique mathématique donnée au cadre interactionnel conformément à sa spécification.

cérité (voir section 3.3) et de faciliter le traitement des aspects sociaux de la communication. Parmi ces approches que nous qualifierons de *conventionnelles*, les jeux de dialogues [REE 98, DAS 00, MAU 01a] semblent émerger comme alternative entre les approches strictement « mentaliste » ou strictement « sociale » qui ne spécifient rien de la structure des dialogues (qui devra « émerger ») et les approches par protocoles qui réduisent l'espace de recherche des continuations possibles au strict minimum, mais font perdre souplesse et adaptativité aux conversations.

Les cadres interactionnels proposés par les approches conventionnelles tels que les jeux de dialogues ne fournissent cependant pas les éléments théoriques nécessaires pour leur utilisation par des agents cognitifs. L'introduction de la couche publique composée d'engagements sociaux nécessite de repenser une pragmatique, une théorie de l'utilisation étendue à ce nouveau cadre. En effet, l'agent ne doit plus raisonner directement sur les intentions (communicationnelles ou pas) des autres, mais sur les engagements pris et à prendre. Ces engagements sont autant les engagements posés ou souhaités par l'agent lui-même que ceux posés par les autres agents, issus des conversations tenues ou des conventions imposées par le cadre interactionnel ou le système.

Dans les sections suivantes, nous étudions plus en profondeur la manière dont un lien pourrait être établi entre les aspects publics (représentés par les engagements) et les aspects privés aux agents (essentiellement leurs états mentaux). Il s'agit en fait de lier (voir section 5) la notion privée d'intention (voir section 2) avec celle publique d'engagement (voir section 3). L'objectif est donc de compléter les approches conventionnelles, en particulier les approches par les jeux de dialogues (voir section 4) en isolant les éléments et mécanismes cognitifs appropriés au traitement de la couche sociale publique. Ce faisant, nous pourrons dégager un certain nombre de problèmes théoriques que la communauté SMA devra investir pour que les nouvelles approches conventionnelles des communications entre agents soient utilisables concrètement (voir section 6).

2. L'intention individuelle[2]

L'intention individuelle est une notion très étudiée en philosophie de l'esprit. Les travaux les plus influents en intelligence artificielle (IA) sont ceux de Searle [SEA 83] et Bratman [BRA 87]. On distingue classiquement : intentionnalité dirigée vers le futur et action intentionnelle. Ces deux notions sont liées puisque l'intention dirigée vers le futur mène généralement à la réalisation d'actions intentionnelles. L'intention qui nous intéresse ici est l'intention dirigée vers le futur.

Il faut se garder de confondre désir et intention. Les désirs représentent des états souhaités du monde, ils peuvent être irréalisables, inaccessibles ou bien contradic-

2. Par le qualificatif « individuelle », on souhaite exclure de notre champ d'investigation les travaux sur les intentions partagées et les intentions collectives. On pense à la notion de We-intention telle que développée par Searle [SEA 90] ou par Tuomela [TUO 88].

toires. La délibération est le processus qui sélectionne les désirs poursuivis sur la base des croyances de l'agent. La délibération concerne la production des buts[3] et a été beaucoup étudiée en intelligence artificielle.

En adjoignant les trois ingrédients que sont croyances, désirs et intention, nous nous retrouvons dans le cadre intentionnel BDI [Belief, Desire, Intention] [RAO 95] classique. Dans la délibération de sens commun (appelée raisonnement pratique), les buts peuvent être interprétés comme des désirs « choisis » suite à une délibération. En outre, les buts doivent être consistants. Finalement, les intentions peuvent être comprises comme des engagements individuels sur des buts, pour lesquels des ressources sont allouées.

Une des caractéristiques principales de l'intention individuelle est donc qu'elle implique un type spécial d'engagement à l'action [BRA 90, WRI 80]. C'est-à-dire que tant que l'agent a l'intention d'atteindre un certain état, il est engagé individuellement à réaliser les actions qu'il pense appropriées pour atteindre cet état. Classiquement, on distingue deux types d'intention, *l'intention de* et *l'intention que*, selon que l'argument est une (ou plusieurs) action ou une proposition [BEL 95, GRO 96]. L'*intention que* s'applique à une proposition (éventuellement complexe) tandis que l'*intention de* est une intention de réaliser une action ou l'intention qu'une action soit réalisée (éventuellement par un autre agent ou un groupe d'agents).

3. Engagements sociaux

La notion d'engagement social ne doit pas être confondue avec la notion d'engagement individuelle évoquée ci-dessus. Conceptuellement, les engagements sociaux capturent les obligations que contractent les agents les uns envers les autres [CAS 95]. En effet, les engagements sociaux sont orientés et indiquent les responsabilités d'un agent envers un autre[4] concernant une action à accomplir ou une proposition à maintenir [SIN 99b]. Les engagements jouent un rôle central dans le dialogue [WAL 95]. Dans la littérature, on leur attache les caractéristiques suivantes (reprises, adaptées et complétées de [MAU 01a]) :

– *Les engagements sont sociaux* : ce sont des engagements vis-à-vis d'autres membres d'une communauté (différent de l'engagement individuel utilisé par Cohen et Levesque [COH 90] pour traduire la persistance liée à l'intention). La particularité principale des engagements sociaux est qu'ils doivent être socialement établis. Finalement, on nomme créditeur l'agent engagé et débiteurs celui ou ceux envers qui l'engagement est pris.

3. Les liens entre les notions de but et d'intention individuelle soulèvent de nombreux débats [HOB 90].
4. Aucun des formalismes rencontrés ne considère l'engagement conjoint de plusieurs agents envers plusieurs autres. Il s'agit d'une première simplification de modélisation.

– *Les engagements sont publics* : les engagements sont publics et accessibles à tous[5]. Pour ce faire, ils sont stockés dans une structure de données commune ou accessible à tous. Dans les systèmes utilisés jusqu'alors, ces structures sont de deux types, l'une est centralisée et l'autre répartie : tableau de conversation ou agendas personnels publics.

– *Les engagements sont propositionnels ou en action* : les engagements propositionnels concernent les éléments propositionnels (par exemple : je crois que p) alors que les engagements en action concernent les actions futures (par exemple : demain, je ferais α).

– *Les engagements peuvent être directs ou conditionnels* : un engagement conditionnel est un engagement qui tient si certaines conditions sont remplies. Les engagements conditionnels sont à rapprocher de l'idée de règle, comme par exemple dans : si tu fais α, je ferais β.

– *Les engagements sont dialogiques ou extra-dialogiques* : parmi les engagements pris au cours du dialogue, on distingue ceux qui concernent le dialogue en cours qui sont dit « dialogiques » de ceux qui se rapportent au contexte du dialogue (la tâche sous-jacente ou le sujet de dialogue). Par exemple, dans les cadres interactionnels conventionnels, lorsqu'une question est posée, l'interlocuteur est engagé à y répondre. Cet engagement dialogique, qui vise à assurer la cohérence structurale du dialogue et qui disparaîtra lorsque le dialogue sera terminé, ne doit pas être confondu avec l'engagement extra-dialogique qui peut découler de la réponse donnée et qui sera généralement persistant au-delà du dialogue.

– *Les engagements sont explicites ou implicites* : les engagements explicites résultent d'actes (de langage ici) alors que les engagements implicites correspondent à des habitudes qui sont incarnées par les règles du cadre interactionnel que les agents sont tenus de respecter ainsi que les conventions et normes en cours dans le système considéré. Les engagements dialogiques sont un exemple d'engagement implicite.

– *Les engagements sont datés* : un engagement est pris à un moment donné et dans un contexte donné[6]. En outre, dans les engagements en action, un des paramètres de l'action doit indiquer à quel moment l'action doit être effectuée (cette indication est donnée sous la forme d'un instant ou d'un intervalle selon le type de gestion temporelle utilisé dans le système considéré).

– *Les engagements sont ordonnés* : certains engagements en action doivent être satisfaits avant d'autres. C'est un moyen classique pour structurer le tableau de conversation (ou les agendas) qui les contient que de les ordonner.

La section suivante, détaille et discute des opérations qui peuvent être appliqués aux engagements sociaux ainsi définis.

---

5. Là encore, il s'agit d'une simplification de modélisation.
6. Par soucis de simplification la notion de contexte n'a pas été prise en compte dans les formalismes rencontrés

3.1. *Opérations sur les engagements sociaux*

En nous inspirant des propositions de Singh [SIN 99a], voyons qu'elles sont les opérations possibles sur les engagements sociaux :

1) *création* : on a vu que la particularité principale des engagements sociaux est qu'ils doivent être socialement établis. Cela signifie qu'un agent ne peut, a priori, pas modifier seul des engagements sociaux. Les opérations sur les engagements sociaux doivent donc être réalisées conjointement pour être valides. Dès lors, un engagement social ne peut être crée que conséquemment à l'adoption d'une convention par une société d'agent ou suite à un dialogue entre agents ;

2) *satisfaction* : un engagement en action est satisfait lorsque l'action dont il est l'objet est exécutée. C'est cette exécution qui génère la satisfaction. En cas de non-satisfaction d'un engagement (l'action n'a pas été effectuée ou encore a été effectuée avec retard. . .On dira que l'engagement n'a pas été *respecté*) dans notre modèle comme dans celui de Maudet et Chaib-draa [CHA 02] des sanctions[7] s'appliquent alors. Ces sanctions sont attachées à l'engagement lors de sa création et ont pu faire l'objet d'une négociation entre les agents. La satisfaction d'un engagement supprime (ou décharge selon les terminologies) l'engagement. Le problème de la satisfaction des engagements propositionnels reste ouvert. A priori, les engagements propositionnels tiennent jusqu'à ce qu'ils soient annulés. Les sanctions attachées à cette annulation dépendent du moment où celle-ci intervient. Un engagement propositionnel qui fait partie d'une argumentation ne peut être annulé sans invalider celle-ci. Aucune proposition fondée n'a été faite à ce sujet, le débat reste ouvert. Un problème subsiste donc à ce niveau pour les engagements propositionnels. Un agent qui s'engage sur p ne pourra plus changer d'avis sans endosser les sanctions attachées à ces anciens engagements propositionnels. En attendant que le problème de satisfaction des engagements propositionnels évoqué ici soit résolu par une solution fondée, on peut supposer que la proposition contient l'intervalle temporel pendant lequel le débiteur est engagé. De la même manière que l'on suppose que l'action dont il est question dans un engagement en action précise l'intervalle temporel dans lequel elle doit être effectuée pour que l'engagement soit satisfait ;

3) *annulation* : l'annulation supprime l'engagement. Que l'annulation provienne du créditeur (on parle alors d'annulation) ou du débiteur (on dit alors que le débiteur renonce à son engagement) des sanctions (éventuellement vides) s'appliquent. Parmi ces sanctions, attachées à l'engagement depuis sa création, on distingue les sanctions attachées au créditeur qui s'appliquent si c'est le créditeur qui annule ou entraîne la non-satisfaction de l'engagement de celles attachées au débiteur qui s'appliquent si il renonce à son engagement ou qu'il est responsable de l'annulation ou de la non-satisfaction de l'engagement ;

4) *La délégation* : le créditeur d'un engagement social peut décider de rendre un autre agent créditeur de celui-ci. Symétriquement, le débiteur peut avoir trouvé un autre débiteur pour un engagement précédemment contracté. Pour des raisons de sim-

7. Ces sanctions peuvent être éventuellement vides.

plification, cette opération n'est pas considérée ici (les systèmes dans lesquelle cette opération est permise restent très flous sur son application).

Dès lors, un engagement sera représenté par un six-uplet $< x, y, c, t, s_x, s_y >$ et noté $C(x, y, c, t, s_x, s_y)$ où x et y sont le débiteur et le créditeur respectivement, c est une proposition ou une action selon qu'il s'agit d'un engagement propositionnel ou d'un engagement en action respectivement, t est la date ou l'engagement a été pris, s_x regroupe les sanctions attachées au débiteur en cas de non-respect de l'engagement ou en cas d'annulation et finalement, s_y est l'ensemble des sanctions pour l'annulation, attachées au créditeur.

Dans les sous-sections suivantes, nous expliquons pourquoi et comment cette notion d'engagement social facilite le traitement des aspects sociaux de la communication entre agents. La sous-section 3.2 indique comment la vérifiabilité sémantique est améliorée et le traitement des aspects sociaux facilité. La sous-section 3.3 indique comment le problème de l'hypothèse de sincérité des approches mentalistes est résolu.

3.2. *Engagements et vérifiabilité et traitements des aspects sociaux.*

Une sémantique basée sur les engagements sociaux qui sont des entités publiques permet de faciliter la vérification de la sémantique des dialogues. En effet, le caractère public des engagements sociaux les rend visibles et accessibles à tous. Ainsi, grâce aux engagements sociaux publics, les agents - comme les concepteurs du système - peuvent vérifier qu'un agent agit conformément à la sémantique des conversations tenues. En effet, suite à une conversation, il n'est plus nécessaire de vérifier si les agents ont modifié leurs intentions et autres états mentaux pour tenir compte de cette conversation, il suffit de vérifier si les engagements correspondants sont présents et on pourra déterminer le moment venu s'ils ont été satisfaits ou non.

Par exemple, à la suite d'une persuasion réussie, les agents ne peuvent vérifier si l'agent persuadé a réellement changé ses états mentaux. Par contre, ils peuvent vérifier que l'engagement propositionnel discuté est établi au niveau de la couche sociale, publique (le tableau de conversation ou les agendas des agents).

La communication entre agents cognitifs se fait par message. Pour éviter de surcharger le système, ces messages sont orientés. Ils ne sont généralement pas transmis par diffusion (broadcast). De ce fait, il est difficile de partager les informations. La couche publique contenant l'historique des engagements pris, nous semble à ce niveau une bonne abstraction pour rendre les communications plus transparentes (lorsque cela est souhaitable) et faciliter le traitement des aspects sociaux dans les SMAs (confiance, réputation, relations sociales. . .). Ce niveau public permet en particulier aux agents de s'observer les uns les autres plus aisément. Cette transparence nous semble essentielle dans des systèmes hétérogènes et ouverts. Par exemple, grâce à cette couche publique, un agent peut raisonner sur la fiabilité d'un autre agent en observant si ce dernier respecte ou non ses engagements.

Finalement, la couche publique permet de garder trace de l'ensemble des interactions, le traitement de la sécurité et des aspects légaux des SMAs s'en trouve facilité (particulièrement s'ils sont ouverts et hétérogènes). Par exemple, si des agents ont à manipuler des informations importantes (valeurs pécuniaire, ...) et dans les cas ou les conséquences de leurs engagements ne sont pas bénignes, les engagements peuvent servir de traces légales. On pense au commerce électronique, où les engagements pourraient être stockés sur des supports sécurisés. En effet, dans le paradigme mentaliste, il est difficile de vérifier (dynamiquement ou rétrospectivement) si un agent a oui ou non promis un paiement à un autre agent. Avec une couche publique mémorisée explicitement, il est aisé de déterminer si un tel engagement a été pris ou pas et s'il a été respecté ou non.

Le problème de savoir si cette promesse était réelle ou bien issue d'une erreur de programmation du concepteur ou de tout autre problème interne à l'agent, reste ouvert, mais les responsabilités sont claires, car les engagements sont transparents. Ce dernier problème est celui de la vérification sémantique informatique[8]. Il consiste à déterminer si un agent implémente correctement les spécifications une sémantique formelle donnée, ou a contrario les règles de fonctionnement d'un cadre interactionnel donné. Il n'est pas résolu par cette approche et relève plutôt du domaine du génie logiciel que des cadres interactionnels eux-mêmes.

3.3. *Engagements et sincérité*

Les engagements sociaux permettent de résoudre le problème de l'hypothèse de sincérité rencontré avec les approches mentalistes. En effet, les engagements sociaux ne sont pas nécessairement sincères. C'est-à-dire que l'on est capable de donner aux primitives de communications une sémantique à base d'engagements, indépendante - a priori - des états mentaux des agents. Ainsi, avec les cadres interactionnels conventionnels, les agents peuvent être insincères. L'insincérité réside dans l'éventuelle différence entre les intentions réelles de l'agent et les engagements poursuivis (qui simulent la réalisation sociale des intentions correspondantes). C'est en fait la théorie de l'utilisation (la pragmatique) des cadres interactionnels ainsi définis qui sera chargée d'établir un lien entre la couche publique constituée d'engagements et les états mentaux privés des agents. Ce lien ne sera cependant pas nécessairement direct et contraignant comme c'est le cas avec les approches mentalistes.

La section suivante présente un langage de communication agents par jeux de dialogue entièrement basé sur la notion de d'engagement social.

8. À contraster à celui de la vérification sémantique « linguistique » discuté jusqu'alors.

4. Les jeux de dialogues de DIAGAL

Les jeux de dialogues sont des structures bilatérales pour la structuration des dialogues. Le modèle que nous avons proposé avec Maudet et Chaib-draa, DIAGAL [DIAlogue Game Agent Language] [CHA 02], se base sur la distinction entre les éléments de dialogue et les éléments de contrôle ou de contextualisation [MAU 01b].

4.1. *Contrôle : acceptation, établissement et structuration*

Dans DIAGAL, le *jeu de contextualisation*, joue le rôle de couche de contrôle. Un métadialogue de contrôle permet aux agents de décider conjointement des jeux de dialogue à jouer et de négocier dynamiquement les différentes structurations de ceux-ci. En particulier, c'est le jeu de contextualisation qui assure l'acceptation (notion essentielle aux interactions dialogiques, dérivée des travaux de Clark [CLA 96] et dont la prise en compte dans les systèmes artificiels a été justifiée par Traum [TRA 94]). Le tableau 1 présente le jeu de contextualisation. Y sont définis, en terme d'engagements dialogiques (indicés par le jeu auquel ils se rapportent, ici j), les coups suivants :

– La proposition d'entrer dans le jeu ($prop.in(x, y, j)$) : propose le jeu j et engage le partenaire à accepter ($acc.in(y, x, j)$), refuser ($ref.in(y, x, j)$) ou (| pour le OU exclusif) proposer un pré-séquencement ($prop.in(y, x, j')$).

– La proposition de fermer un jeu ($prop.out(x, y, j)$) : propose la fermeture du jeu j et engage l'interlocuteur à accepter ou bien refuser la proposition.

– L'acceptation de fermeture d'un jeu ($acc.out(x, y, j)$) : supprime tous les éventuels engagements dialogiques liés au jeu j qui n'ont pas encore été satisfait[9].

– l'acceptation d'ouverture d'un jeu ($acc.in(x, y, j)$) : indique à l'initiateur du jeu de dialogue que le jeu va être joué. Cette acceptation créée les engagements dialogiques qui correspondent aux règles du jeu, cette dernière notion est présentée dans la section suivante.

– le refus d'une proposition d'entrée ($ref.in(x, y, j)$) : sans effet sur la couche publique, il signifie le refus à l'interlocuteur.

– le refus d'une proposition de sortie ($ref.out(x, y, j)$) : sans effet sur la couche publique, il signifie le refus à l'interlocuteur.

Durant un jeu de dialogue, un jeu peut être imbriqué à tout moment par simple proposition dans le jeu de contextualisation de DIAGAL. Avec ce jeu de contextualisation, trois types de structurations sont donc possibles :

– *le séquencement* : il y a séquencement de deux jeux j_1 et j_2 si le jeu j_2 est proposé alors que le jeu j_1 a été fermé ;

9. On note que les deux agents impliqués dans le dialogue doivent être d'accord pour ouvrir ou fermer un jeu, ce qui est une hypothèse forte et discutable, au moins pour la fermeture.

– *le pré-séquencement* : il y a pré-séquencement lorsque qu'un jeu est ouvert alors qu'un autre est proposé (mais pas encore accepté ou refusé). Le pré-séquencement sert essentiellement à établir les conditions d'entrée dans le jeu pré-séquencé ;

– *l'emboîtement ou imbrication* : un jeu j_1 est emboîté dans un autre j_2 s'il est ouvert alors que j_2 est déjà ouvert. Ainsi, les emboîtements sont gérés via une pile de jeux ouverts. Les engagements du jeu courant sont prioritaires sur les engagements des autres jeux.

Coups	Opérations
$prop.in(x, y, j)$	$create(y, C_j(y, x, acc.in(y, x, j) \| ref.in(y, x, j) \| prop.in(y, x, j')))$
$prop.out(x, y, j)$	$create(y, C_j(y, x, acc.out(y, x, j) \| ref.out(y, x, j)))$
$acc.in(x, y, j)$	créé les engagements dialogiques correspondants au jeu j
$acc.out(x, y, j)$	supprime les engagements dialogiques pendant du jeu j
$ref.in(x, y, j)$	sans effet sur la couche publique
$ref.out(x, y, j)$	sans effet sur la couche publique

Figure 1. *Jeu de contextualisation de DIAGAL.*

4.2. *Quatre jeux de dialogue primitifs*

Concrètement, chaque jeu de dialogue DIAGAL est défini à la façon des systèmes dialectiques par :

1) *des conditions d'entrée*, E, qui doivent être respectées pour que le jeu puisse être proposé ;

2) *des conditions de succès* pour l'initiateur, SI, qui sont remplies si le jeu est un succès en faveur de l'initiateur, c'est-à-dire que l'engagement recherché est accepté ;

3) *des conditions d'échec* pour l'initiateur, FI, qui sont remplies si l'engagement recherché est rejeté à l'issue du jeu ;

4) *un ensemble de règles* R qui sont exprimées en terme d'engagements dialogiques éventuellement conditionnels et en terme d'engagements extra-dialogiques (sans indice).

Comme pour le jeu de contextualisation, toutes ces notions sont définies en terme d'engagements, éventuellement conditionnels. On y distinguer les engagements dialogiques, qui sont issus des règles du jeu de dialogue, indicés j et les engagements extra-dialogiques (non indicés). Lorsqu'un jeu est fermé, les engagements dialogiques qui en sont issus et qui n'ont pas été satisfaits pendant le jeu sont supprimés, mais pas les engagements extra-dialogiques qui correspondent aux conditions de succès ou d'échec si celles-ci ont été atteintes.

Un jeu de dialogue a pour fin, s'il se déroule avec succès (c'est-à-dire que les conditions de succès sont atteintes), de poser, créer, *un* nouvel engagement extra-dialogique. Or, il existe quatre types d'engagements directs[10] qu'un locuteur x qui prend l'initiative d'un jeu de dialogue avec un interlocuteur y peut envisager obtenir :

1) un engagement propositionnel du locuteur envers l'interlocuteur : $C(x, y, p, t, s_x, s_y)$ $(inform)$

2) un engagement propositionnel de l'interlocuteur envers le locuteur : $C(y, x, p, t, s_x, s_y)$ (ask)

3) un engagement en action du locuteur envers l'interlocuteur : $C(x, y, \alpha, t, s_x, s_y)$ $(offer)$

4) un engagement en action de l'interlocuteur envers le locuteur : $C(y, x, \alpha, t, s_x, s_y)$ $(request)$

Dans DIAGAL, un jeu de dialogue est défini pour chacun de ces types d'engagement (le nom en est indiqué entre parenthèses dans la liste ci-dessus). Les figures 2, 3, 4 et 5 présente le quadruplet de définition pour les jeux $request$, $offer$, $inform$ et ask sus-nommés, à savoir $< E_r, SI_r, FI_r, R_r >$, $< E_o, SI_o, FI_o, R_o >$, $< E_i, SI_i, FI_i, R_i >$ et $< E_a, SI_a, FI_a, R_a >$ respectivement.

$$
\begin{array}{l|l}
E_r & \neg C(y, x, \alpha, t_0) \\
SI_r & C(y, x, \alpha, t_f) \\
FI_r & \neg C(y, x, \alpha, t_f) \\
R_r & C_j(x, y, \mathit{request}(x, y, \alpha), t_0) \\
& C_j(y, x, \mathit{request}(x, y, \alpha) \Rightarrow \\
& C_j(y, x, \mathit{promise}(y, x, \alpha) | \mathit{refuse}(y, x, \alpha), t_1), t_0) \\
& C_j(y, x, \mathit{promise}(y, x, \alpha) \Rightarrow C(y, x, \alpha, t_2), t_0) \\
& C_j(y, x, \mathit{refuse}(y, x, \alpha) \Rightarrow \neg C(y, x, \alpha, t_2), t_0)
\end{array}
$$

Figure 2. *Conditions et règles du jeu $request$.*

$$
\begin{array}{l|l}
E_o & \neg C(x, y, \alpha, t_0) \\
SI_o & C(x, y, \alpha, t_f) \\
FI_o & \neg C(x, y, \alpha, t_f) \\
R_o & C_j(x, y, \mathit{offer}(x, y, \alpha), t_0) \\
& C_j(y, x, \mathit{offer}(x, y, \alpha) \Rightarrow \\
& C_j(y, x, \mathit{accept}(y, x, \alpha) | \mathit{refuse}(y, x, \alpha), t_1), t_0) \\
& C_j(x, y, \mathit{accept}(y, x, \alpha) \Rightarrow C(x, y, \alpha, t_2), t_0) \\
& C_j(x, y, \mathit{refuse}(y, x, \alpha) \Rightarrow \neg C(x, y, \alpha, t_2), t_0)
\end{array}
$$

Figure 3. *Conditions et règles du jeu $offer$*

Ces quatre jeux de base forment un système d'opérateurs « complet » pour l'ajout d'engagements extra-dialogiques directs dans la couche publique. En effet, les quatre

10. Par simplification, on considère la création d'engagement extra-dialogiques conditionnels comme un problème théorique ouvert (section 6.

$$
\begin{array}{l|l}
E_i & \neg C(x,y,p,t_0) \text{ and } \neg C(x,y,\neg p,t_0) \\
SI_i & C(x,y,p,t_f) \\
FI_i & \neg C(x,y,p,t_f) \\
R_i & C_j(x,y,assert(x,y,p),t_0) \\
 & C_j(y,x,assert(x,y,p) \Rightarrow \\
 & C_j(y,x,agree(y,x,p)|disagree(y,x,\neg p),t_1),t_0) \\
 & C_j(x,y,agree(y,x,p) \Rightarrow C(x,y,p,t_2),t_0)
\end{array}
$$

Figure 4. *Conditions et règles du jeu* $inform$

$$
\begin{array}{l|l}
E_a & \neg C(y,x,p,t_0) \\
SI_a & C(y,x,p,t_f) \\
FI_a & \neg C(y,x,p,t_f) \\
R_a & C_j(x,y,ask(x,y,p),t_0) \\
 & C_j(y,x,ask(x,y,p) \Rightarrow \\
 & C_j(y,x,agree(y,x,p)|disagree(y,x,p),t_1),t_0) \\
 & C_j(x,y,agree(y,x,p) \Rightarrow C(y,x,p,t_2),t_f) \\
 & C_j(x,y,disagree(y,x,p) \Rightarrow \neg C(y,x,p,t_2),t_f)
\end{array}
$$

Figure 5. *Conditions et règles du jeu* ask

types d'engagements que permettent de poursuivre les jeux de DIAGAL sont les seuls envisageables selon les typologies actuelles. C'est important de le noter, car ce n'est pas sans conséquence. Cela signifie que toutes les modifications de la couche publique qui ont la forme d'un ajout d'engagement peuvent être obtenues par utilisation de ces quatre jeux. Ce résultat tient sous les hypothèses simplificatrices suivantes : (1) le dialogue mène à l'ajout d'engagements directs et non d'engagements conditionnels et (2) les engagements sont posés séquentiellement, un par un.

Même si aucun jeu n'est défini pour le retrait d'engagement, notons que le système n'est pas « monotone »[11] car les engagements peuvent être retirés, supprimés, en étant : satisfait, non respectés ou annulés. Dans les deux derniers cas, des sanctions (attachées à l'engagement depuis sa création et ayant pu faire l'objet de négociation) sont appliquées pour l'agent responsable.

5. Lier les cognitions publiques aux cognitions privées

Un agent insatisfait par une situation donnée peut agir selon deux modes bien distincts. Il peu essayer de modifier son environnement pour que celui-ci corresponde à la situation attendue (c'est-à-dire faire des tentatives pour réaliser ses intentions).

11. Il n'a pas été posé explicitement comme hypothèse de modélisation que les jeux de dialogue ne devrait être conçus que pour l'ajout d'engagements.

Il peut tout autant modifier ses états mentaux pour que ceux-ci correspondent à la situation extérieure (c'est-à-dire changer ses états mentaux et dans notre cas, renoncer à certaines intentions, quitte à en adopter d'autres).

Puisque les agents raisonnent sur la base de leurs états mentaux privés, mais à propos d'engagements publics. Il est nécessaire d'établir un lien entre ces types de cognitions. Pour ce faire, on considère qu'*un engagement en action est le pendant socialement accepté d'une « intention de » et un engagement propositionnel est le pendant socialement accepté d'une « intention que »*.

Dans les SMAs, la question de savoir quand l'agent doit essayer de modifier l'environnement (entre autres, la couche publique des engagements sociaux) pour satisfaire ses intentions, et quand l'agent doit modifier ses états mentaux pour être cohérent avec son environnement est cruciale.

Les engagements sociaux sont des cognitions particulières qui ne sont pas modifiables individuellement, mais doivent être socialement établis (c'est l'objet des jeux de dialogue que de fournir des outils pour établir des engagements conjointement acceptés). C'est-à-dire que pour créer un engagement, un agent doit dialoguer. C'est par le dialogue que les agents vont essayer de faire établir les engagements sociaux cohérents avec leurs autres cognitions, en particulier les engagements correspondant à leurs intentions. Cependant, à l'issue de ces dialogues, certains engagements peuvent rester incohérents tout en n'étant plus modifiables, car on a vu que les engagements posés font office d'obligations sociales. Pour réduire d'éventuelles incohérences tout en se conformant aux engagements pris, c'est alors ses états mentaux que l'agent devra changer pour rétablir la cohérence. C'est le ressort du changement d'attitude et cela formalise la vision des psychologues Brehm et Cohen à ce sujet [BRE 62], soutenue par de nombreuses d'expériences.

C'est-à-dire que par le dialogue pour chaque « intention de », l'agent va chercher à obtenir l'engagement propositionnel correspondant et pour chaque « intention que », l'engagement en action correspondant. Les dialogues ne sont cependant que des tentatives dans ce sens, qui peuvent échouer. Les jeux de dialogue fournissent un bon outil pour structurer et réaliser ces tentatives.

La figure 6 résume ce nouveau paradigme dans lequel les agents ne raisonnent plus uniquement sur leurs intentions et celles des autres, mais sur les engagements pris et à prendre. On y présente de gauche à droite : les intentions issues du raisonnement sur les autres cognitions (dont certaines sont privées, comme les désirs et les croyances, tandis que d'autres sont publiques, comme les engagements sociaux), les engagements correspondant à ces intentions (selon les liens établis dans cette section) et finalement les jeux de dialogue DIAGAL qui doivent être utilisés pour tenter de faire accepter l'engagement souhaité.

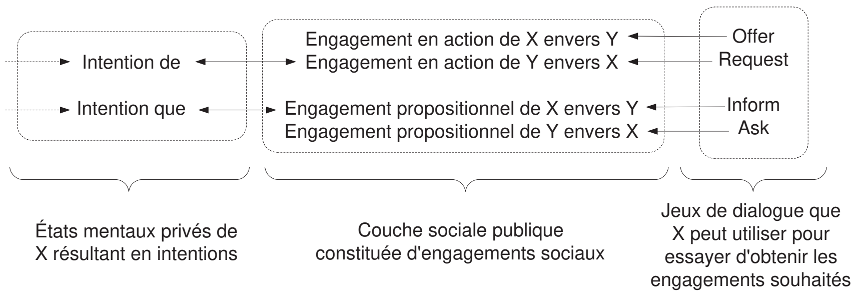

Figure 6. *Liens entre les cognitions publiques, privées et les jeux de dialogue.*

5.1. *Utiliser les jeux de dialogue*

À partir des liens esquissés dans la section précédente, on peut tracer les grandes lignes d'un schéma général pour l'utilisation dynamique des jeux de dialogues définis dans DIAGAL :

1) l'agent produit ses intentions via le raisonnement pratique (sur la base de ces croyances, désirs et buts) ;

2) l'agent sélectionne l'intention courante ;

3) si nécessaire, l'agent agit pour satisfaire cette intention ;

4) si nécessaire, l'agent communique pour essayer d'obtenir un engagement qui serait le pendant socialement accepté de son intention.

5) si la communication est un succès en sa faveur, l'engagement poursuivi tient (retour étape 1). Des sanctions tiendront si un engagement en action pris n'est pas respecté ou si un engagement propositionnel est contredit et s'avère erroné.

6) si la communication échoue, l'agent doit se conformer aux engagements qui en sont issus et réviser ses états mentaux en fonction de ceux-ci, afin d'être cohérent avec le niveau social.

Par exemple, si un agent A a l'intention qu'un agent B effectue l'action α, il doit essayer d'obtenir l'engagement correspondant à son « intention que ». Cet engagement est de la forme $C(B, A, \alpha, t, s_B, s_A)$. Le jeu de dialogue DIAGAL permettant de réaliser cette tentative est le jeu $request$ comme l'indique la figure 6. En effet, seule la condition de succès du jeu $request$ est un engagement en action de la forme désirée. Si la communication est un succès en sa faveur, l'engagement désiré sera créé dans la couche publique et B devra le satisfaire sous peine d'écoper des sanctions attachées au non-respect dans s_B.

6. Conclusion et problèmes théoriques ouverts

Il reste un certain nombre de problèmes qui ne sont traités par aucune des approches conventionnelles rencontrées jusqu'alors. Parmi ceux-ci, on trouve :

– *l'élicitation du modèle d'autrui au profit de la couche publique* : le raisonnement sur la couche publique sera-t-il suffisant pour permettre tous les traitements sociaux ? En particulier, dispense-t-il du maintien et du traitement d'un modèle d'autrui ? Avec les cadres interactionnels conventionnels, les agents peuvent être insincère. La non-sincérité réside dans l'éventuelle différence entre les intentions réelles de l'agent et les engagements poursuivis (qui simulent les intentions correspondantes). Pour se représenter cette insincérité, il est donc nécessaire de maintenir un modèle d'autrui. Ainsi, sans modèle d'autrui, les agents ne pourront se représenter la non-sincérité de cette manière, mais peut-être pourront-ils l'inférer du non-respect de certains engagements. Pour l'heure et en l'état des modèles actuels, la question reste ouverte aussi bien en pratique qu'en théorie (on a là deux définitions différentes de la sincérité).

– *la gestion de la cohérence de la couche publique* : dans les approches conventionnelles, on ne voit pas ce qui peut garantir que les engagements pris sont cohérents (même si les conditions d'entrée des jeux DIAGAL limitent les incohérences possibles). Il semble que les dépendances entre engagements ne soient pas prises en compte. Par exemple, un agent A a initié une requête envers un agent B pour qu'il ferme la porte. Si un sous-jeu d'offre est joué par B avec succès pour garder la porte ouverte avant qu'une réponse ne soit donnée dans le jeu principal (de sorte que les conditions d'entrée dans le sous-jeu d'offre soient vérifiées), alors des engagements incohérents (en l'occurrence, $C(B, A, t, F, s_B, s_A)$ et $C(B, A, t, \neg F, s_B, s_A)$ où F décrit l'action de fermer la porte) peuvent résulter du dialogue.

– *la définition et la mesure de l'utilité des dialogues* : dans les approches par jeux de dialogue rencontrées, rien ne garantit que les conversations tenues soient utiles aux agents en terme de leurs objectifs individuels ou collectifs ;

– *la terminaison des dialogues* : toutes les approches par les jeux de dialogue rencontrées permettent la composition de jeux par imbrication. Cependant, rien n'indique si cette possibilité de composition est limitée. De ce fait, rien ne garantit la fin des dialogues (c'est-à-dire que rien ne garantit que les jeux ouverts soient fermés dans un temps fini) ;

– *la puissance expressive du cadre interactionnel* : aucune des approches rencontrées jusqu'alors ne permet de créer des engagements sociaux extra-dialogiques conditionnels. Une telle possibilité permettrait pourtant d'accroître la puissance expressive des approches conventionnelles. On peut considérer que les conditions sont exprimées dans le contenu propositionnel ou l'action véhiculée par l'engagement extra-dialogique négocié dans les jeux de dialogue DIAGAL existants. Sinon, on peut imaginer un nouveau type de composition dans le jeu de contextualisation spécifiquement dédié à la création d'engagements extra-dialogiques conditionnels.

7. Conclusion

Dans cet article, nous avons présenté le nouveau paradigme de communication agents qui nous semble sous-jacent (et généralement implicite) aux approches conventionnelles récemment proposées pour la communication agents. En effet, ces approches, caractérisées par une meilleure prise en compte des aspects sociaux des communications, s'articulent autour de la notion d'engagement social et impliquent le changement de paradigme suivant :

> Les agents cognitifs ne doivent plus simplement raisonner sur leurs intentions et celles des autres, mais sur les engagements sociaux pris et à prendre.

Tout en présentant un cadre interactionnel complet (le cadre DIAGAL, voir section 4) pour la manipulation de la couche publique qui contient les engagements sociaux, nous avons tenté d'établir les liens nécessaires pour pouvoir tracer les grandes lignes de son utilisation par des agents cognitifs. Ce faisant, nous avons essayé d'expliciter certains avantages (voir entre autres les sections 3.2 et 3.3) faiblesses de ces nouvelles approches.

À la lumière de ces considérations, il apparaît que la définition d'un cadre interactionnel conventionnel basé sur les jeux de dialogue ne donne pas toutes les informations nécessaires quant à son utilisation. C'est, évidemment, la cognition de l'agent qui devra prendre en charge bon nombre de ces aspects. Autrement dis, le changement de paradigme annoncé ici devra s'accompagner de travaux sur les modèles cognitifs et les théories d'agents de façon à y intégrer le raisonnement sur la couche publique. Pour ce faire, il est nécessaire d'investir de nouveau les aspects pragmatiques de la communication dans les théories d'agents. C'est ce que nous avons souhaité montrer ici tout en offrant quelques éléments et pistes de solutions (section 5)[12].

8. Bibliographie

[AMG 00] AMGOUD L., MAUDET N., PARSON S., « Modelling Dialogue Using Argumentation. », *Proceedings of the 4th internationnal conference on multi-agent systems (ICMAS'00)*, 2000.

[BEL 95] BELL J., « Changing Attitudes », WOOLDRIDGE M., JENNINGS N. R., Eds., *Intelligent Agents : Theories, Architectures, and Languages (LNAI Volume 890)*, Springer-Verlag : Heidelberg, Germany, janvier 1995, p. 40–55.

[BRA 87] BRATMAN M. E., *Intentions, Plans, and Practical Reason*, Harvard University Press : Cambridge, MA, 1987.

[BRA 90] BRATMAN M. E., « What is Intention ? », COHEN P. R., MORGAN J. L., POLLACK M. E., Eds., *Intentions in Communication*, p. 15–32, The MIT Press : Cambridge, MA, 1990.

12. Notons que nous avons proposé, par ailleurs, une pragmatique des communications agents basée sur la notion de cohérence cognitive [PAS 02].

[BRE 62] BREHM J., COHEN A., *Explorations in Cognitive Dissonance*, John Wiley and Sons, inc, 1962.

[CAS 95] CASTELFRANCHI C., « Commitments : from Individual Intentions to Groups and Organizations », *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, San Francisco, CA, juin 1995, p. 41–48.

[CHA 02] CHAIB-DRAA B., MAUDET N., LABRIE M., « Request for Action Reconsidered as Dialogue Game based on Commitments », *Workshop on Agent Communication Language (AAMAS´02)*, 2002.

[CLA 96] CLARK H. H., *Using Language*, Cambridge University Press, 1996.

[COH 90] COHEN P. R., LEVESQUE H. J., « Intention is Choice with Commitment », *Artificial Intelligence*, vol. 42, 1990, p. 213–261.

[COL 98] COLOMBETTI M., « Different ways to have something in common », CHRISTIANSEN H., ANDREASEN T., LARSEN H. L., Eds., *Procedings of the third internationnal conference on flexible query answering systems*, Springer, Berlin, 1998, p. 95-109.

[COL 00] COLOMBETTI M., « Commitment-Based Semantic for Agent Communication Languages », *1st Workshop on the History and Philosophy of Logic, Mathematics and Computation*, 2000.

[DAS 00] DASTANI M., HULSTIJN J., DER TORRE L., « Negociation Protocols and Dialogue Games », *proceedings of the BNAIC*, 2000.

[DIG 00] DIGNUM F., « Issues in Agent Communication : An Introduction », DIGNUM F., GREAVES M., Eds., *Issues in Agent Communication*, Lecture Notes in Artificial Intelligence, Springer-Verlag : Heidelberg, Germany, 2000, p. 1-16.

[FIN 94] FININ T., FRITZSON R., « Specification of the KQML Agent Communication Language plus Example Agent Policies and Architectures », rapport, 1994, DARPA Knowledge Sharing Initiative External Interface Working Group.

[FIN 98] FININ T., LABROU Y., « Semantics for an ACL », *Agents, Architecture and Languages*, Springer-Verlag : Heidelberg, Germany, Berlin, 1998.

[FIP 02] FIPA, « FIPA [Fondation for Intelligent Physical Agents] », http ://www.FIPA.org, 2002.

[FLO 01] FLORES R., KREMER R., « Bringing Coherence to Agent Conversation », WOOLDRIDGE M., CIANCARINI P., WEISS G., Eds., *Agent-Oriented Software Engineering II*, Lecture Notes in Computer Science, Springer-Verlag, 2001, p. 50-67.

[GRO 96] GROSZ B. J., KRAUS S., « Collaborative Plans for Complex Group Action », *Artificial Intelligence*, , n° 86, 1996, p. 269-357.

[HOB 90] HOBBS J., « *Intentions in communication* », chapitre Artificial Intelligence and Collective Intentionnality, p. 445-459, MIT Press, 1990.

[MAU 01a] MAUDET N., « Modéliser les Conventions des Interactions Langagières : la Contribution des Jeux de Dialogue », PhD thesis, ENSEIHT-UPS, 2001.

[MAU 01b] MAUDET N., CHAIB-DRAA B., « Commitment-based and Dialogue-game based Protocols–News Trends in Agent Communication Language », *Knowledge Engineering*, , 2001.

[MOU 97] MOULIN B., « *Agent and Multi-agent Systems* », chapitre The social dimension of interacions in multi-agent systems, LNAI 1441, Springer, Berlin, 1997.

[PAS 02] PASQUIER P., CHAIB-DRAA B., « Cohérence et conversations entre agents : vers un modèle basé sur la consonnance cognitive », MÜLLER J., MATHIEU P., Eds., *Systèmes multi-agents et systèmes complexes, actes des JFIADSMA'02*, Paris, 2002, Hermes Science Publication, p. 188-203.

[RAO 95] RAO A. S., GEORGEFF M., « BDI Agents : from theory to practice », *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, San Francisco, CA, juin 1995, p. 312–319.

[REE 98] REED C., « Dialogue frames in agent communication », *Proceedings of the Third Internationnal Conference on MultiAgent Systems (ICMAS)*, 1998.

[SEA 69] SEARLE J. R., *Speech Acts : An Essay in the Philosophy of Language*, Cambridge University Press : Cambridge, England, 1969.

[SEA 83] SEARLE J. R., *Intentionality : An Essay in the Philosophy of Mind*, Cambridge University Press : Cambridge, England, 1983.

[SEA 90] SEARLE J. R., « Collective Intentions and Actions », COHEN P. R., MORGAN J., POLLACK M. E., Eds., *Intentions in Communication*, p. 401–416, The MIT Press : Cambridge, MA, 1990.

[SIN 98] SINGH M. P., « Agent Communication languages : rethinking the principles », *IEEE Computer*, vol. 12, n° 31, 1998, p. 40-47.

[SIN 99a] SINGH M., « An Ontology for Commitments in Multiagent Systems : Toward a Unification of Normative Concepts », *Artificial Intelligence and Law*, vol. 7, 1999, p. 97-113.

[SIN 99b] SINGH M., RAO S., GEORGEFF M., « *Multiagent Systems : A Modern Approach to Distributed Artificial Intelligence* », chapitre Formal Methods in DAI : Logic-Based Representation and Reasoning, p. 331-376, The MIT Press, Cambridge, MA, 1999.

[SIN 00] SINGH M., « A Social Semantics for Agent Communication Languages », DIGNUM F., GREAVES M., Eds., *Issues in Agent Communication*, p. 31–45, Springer-Verlag : Heidelberg, Germany, 2000.

[TRA 94] TRAUM D. R., « A Computationnal Theory of Grounding in Natural Language Conversation », PhD thesis, Departement of Computer Sciences, University of Rochester, 1994.

[TUO 88] TUOMELA R., MILLER K., « We-Intentions », *Philosophical Studies*, vol. 53, 1988, p. 367–389.

[WAL 95] WALTON D., KRABBE E., *Commitment in Dialogue*, Suny Press, 1995.

[WRI 80] VON WRIGHT G., *Freedom and determination*, North Holland Publishing Co., 1980.