Institucions Electròniques

Carles Sierra, Pablo Noriega
Institut d'Investigació en Intel·ligència Artificial, IIIA
Consell Superior d'Investigacions Científiques, CSIC
08193 Bellaterra, Barcelona.
sierra@iiia.csic.es

Abstract

En aquest treball proposem la noció d'institució electrònica intermediada per agents com l'estructura social adient per a una gran varietat de sistemes multiagent. Presentem una proposta de formalització d'aquestes organitzacions socials i esmentem tres projectes, de caire ben diferent, en els quals s'està aplicant aquesta noció de manera exitosa.

1 Introducció

La noció d'interacció entre agents forma el nucli central de tots els sistemes multiagents. Les interaccions són la part observable dels diferents mecanismes de comportament d'aquests sistemes: cooperació, coordinació, col·laboració i negociació. En un sentit molt ampli, i degut a la seva generalitat, ens referirem a la negociació en aquest treball com el mecanisme bàsic d'interacció.

El nostre objectiu és construir agents que ajudin els humans a negociar. Els humans poden, eventualment, delegar completament en agents autònoms les tasques de negociació, però, en qualsevol cas, els agents que es desenvolupin han d'estar preparats per a interactuar, no solament amb altres agents autònoms, sinó també amb humans. Per tant, els models d'interacció que ens interessen són, precisament, aquells que tenen en compte les peculiaritats de la negociació humana. En particular, i com a tret fonamental, la noció de *diàleg*. Considerem que la negociació, i la interacció en general, és una activitat dialògica. Els humans i els agents, quan cerquen acords, dialoguen. Es precisament per aquest fet (que els agents poden estar interactuant amb un humà) que no podrem assumir, en general, nocions d'omnisciència o racionalitat perfecta en les contraparts dels nostres agents.

Aquests intercanvis dialògics pressuposen que els agents són entitats capaces d'establir compromisos. En termes de negociació, aquests compromisos vinculen els participants del diàleg gràcies a convencions socials i institucions que els fan acomplir. Per exemple, un acord de divorci s'estableix mitjançant un contracte que és validat i és mediat per un jutge; les transmisions patrimonials requereixen d'un registre civil i de la intervenció de notaris; les accions són intercanviades a institucions —borses.

Noteu també que si bé la negociació humana pot aprofitar-se de recursos elucidatoris i explicatius de la conversa, les negociacions intermediades en les institucions esmentades a dalt obvien aquesta necessitat imposant una interpretació estricta del llenguatge utilitzat en el diàleg a través de convencions establertes per les institucions. Aquest és el cas, per exemple, de les crides de preus d'una casa de subhastes, o de les seqüències de números en una taula de borsa. De la mateixa manera, en les interaccions mediades per agents es necessita fer explícit el significat i, encara que és imaginable que les interaccions dialògiques puguin tenir un contingut elucidatori, resulta molt més pràctic restringir la negociació, i en general les interaccions, a aquelles que es realitzen utilitzant llenguatges perfectament definits. Aquest article es concentra, precisament, en les primeres passes vers la formalització d'institucions electròniques.

La noció d'institució electrònica es fonamenta en tres pilars que contenen els elements esmentats fins ara:

1. Un marc dialògic. Alguns aspectes d'una institució són estables i constitueixen el context o marc de referència de la interacció entre els agents. En una institució dialògica els agents interactuen mitjançant il·locucions.

Les institucions defineixen quines són les il·locucions acceptables, quina és l'ontologia —incloent-hi rols, llocs i temps—, un llenguatge comú per referir-se al "món", un llenguatge comú per comunicar-se, i un metallenguatge comú per a, per exemple, establir preferències. En molts casos, una institució pot considerar com a rellevants relacions socials entre els agents participants —en termes d'autoritat, per exemple—, o alguna mena de característica de "personalitat" que pot afectar les interaccions. Tots aquests aspectes contextuals són el que anomenem marc dialògic.

2. Una estructura performativa. Les interaccions dels agents estan articulades al voltant de trobades de grups d'agents, que anomenem escenes, amb un protocol de comunicació ben definit. Entenem per protocol d'una escena els possibles diàlegs que els agents poden tenir-hi.

3. Unes regles de comportament. Les accions dels agents en el contexte d'una institució, en el cas que ens ocupa, accions dialògiques, tenen unes conseqüències, normalment en forma de compromisos, que imposen restriccions sobre les accions dialògiques dels agents en les escenes en que participaran en el futur. Les regles de comportament afecten el comportament dels agents, delimitant-lo.

La intuició darrera de la noció d'institució és que és un mecanisme molt útil en l'especificació de sistemes d'agents autònoms interactius. A la secció 2 presentem el primer esboç de formalització de la noció d'institució electrònica, i a la secció 3 esmentem tres projectes en els que preveiem l'ús d'aquest concepte.

2 Una proposta formal

Com hem dit a l'apartat anterior, la noció d'institució electrònica es basa en tres pilars ben diferenciats. Definirem primer aquests tres pilars i després les institucions electròniques com a un agrupament dels tres.

Adaptant les definicions de [2, 8] tenim que la noció bàsica per a definir els elements ontològics és el marc dialògic:

Definition 1 *Un* Marc dialògic *és una tupla* $DF = (Rols, L, ML, CL, Time)$*, on*

1. *Rols és un conjunt d'identificadors de rol.*
2. *L és un llenguatge lògic per a representar el domini.*
3. *ML és un metallenguatge sobre L.*
4. *CL és el llenguatge de comunicació entre agents. Es defineix com* $CL = \{\iota(a, b, \varphi, t) | a, b$ *són variables sobre Rols,* ι *és una il·locució,* $\varphi \in L, t \in Time\}$
5. *Time és un conjunt discret i totalment ordenat d'instants.*

La noció d'estructura performativa és la més complexa i interessant del formalisme perquè és l'encarregada de modelitzar la dinàmica del diàleg. La definició es basa en la noció d'escena que presentem a continuació. Tota escena té un estat inicial on comença l'intercanvi d'il·locucions entre els participants en l'escena. En cada moment l'escena es troba en un estat del qual es surt per les transicions que van cap a nous estats. Una transició es realitza si la il·locució que l'etiqueta es produeix. Una escena es considera acabada quan es troba en un dels possibles estats terminals. El conjunt de transicions i estats determina quin és el protocol de l'escena, i.e. el conjunt d'interaccions dialògiques possibles.

Definition 2 *Donat un marc dialògic* $DF = (Rols, L, ML, CL, Time)$*, una escena és una tupla* $s = (A, R, W, w_0, W_f, \rho)$*, on*

A és un conjunt finit de variables d'agent.

$R : A \rightarrow Rols$ assigna un rol a cada variable d'agent en l'escena.

W és un conjunt d'estats. $w_0 \in W$ és l'estat inicial de l'escena i $W_f \subseteq W$ els estats finals de l'escena.

$\rho : W \times CL \rightarrow W$ és una funció de transició entre estats, on les transicions estan etiquetades amb una il·locució.

Una escena es pot "realitzar" quant tenim una assignació d'agents "concrets" (i.e. programes o humans) per a totes i cadascunes de les variable d'A. Un cop l'assignació és completa l'escena es pot començar a executar, començant per l'estat inicial, i fins a acabar en un dels estats finals. Un cop una escena ha finalitzat, els agents que hi han participat es redistribuiran, com veurem en la definició d'estructura performativa. Alguns poden, però, quedar-se en l'escena per tornar a executar-la amb agents nous.

Una estructura social complexa, com és el cas d'una institució, consta de multitut d'escenes interrelaciones que s'executen de forma concurrent o seqüencial segons sigui el cas. El punt fonamental, per tal d'establir la seva interrelació, és determinar quins agents abandonen quina escena, en quin estat l'abandonen, i a quina escena s'incorporen. Les restriccions que s'imposin a aquest fluxe d'agents determinarà en bona mesura la flexibilitat del formalisme. Hi ha una restricció necessària des del nostre punt de vista, i és la integritat dels agents que participen en una escena. Es a dir, per a produir-se una transició entre estats d'una escena donada, tots els agents de l'escena han de ser-hi presents. Dit d'una altra manera, si algun agent abandona una escena, aquesta resta congelada —no s'emeten il·locucions— fins que l'agent s'hi reincorpora. Amb aquesta restricció en ment, podem modelitzat la noció de transició entre escenes d'una manera similar a com es modelitzen les transicions en xarxes de Petri: amb una funció d'incidència previa i una d'incidència posterior. El que si permet el formalisme s que un agent concret participi en més d'una escena alhora. Això ens dóna la flexibilitat necessària per modelitzar el fet que un agent pot participar en més d'un dialeg alhora.

Cal esmentar que la noció d'escena representa el concepte d'especificació d'un conjunt d'agents que segueixen un protocol de dialeg concret. Ara be, en un moment determinat una institució pot permetre l'execució simultània de més d'una particularització d'una escena amb actors diferents.

Notarem amb subindex S la unió per a totes les escenes del conjunt S del diferents components, per exemple, $W_S = \{w | (A, R, W, w_0, W_f, \rho) \in S \text{ and } w \in W\}$.

Definition 3
Donat un marc dialògic $DF = (Rols, L, ML, CL, Time)$ una estructura performativa és una tupla $PS = (S, T, \alpha, \beta)$

S és un conjunt finit i no buit d'escenes.

T és un conjunt finit i no buit de transicions entre escenes.

$\alpha : W_S \times T \rightarrow 2^{A_S}$ és la funció d'incidència prèvia,

$\beta : T \times W_S \rightarrow 2^{A_S}$ és la funció d'incidència posterior.

verificant que

1. *hi ha integritat dels agents: per a tota $t \in T$, $a \in \alpha(w, t)$ per algun $w \in W_S$ si i només si $a \in \beta(t, w')$ per algun $w' \in W_S$.*
2. *hi ha un únic estat actiu en una escena: per a tota $t \in T, w, w' \in W_{s'}$ tals que $w \neq w'$, si $\beta(t, w) \neq \emptyset$ and $\beta(t, w') \neq \emptyset$ llavors $s \neq s'$.*

Entenem que els actes de parla d'un agent introdueixin compromisos d'actuació posteriors. Els compromisos s'han d'interpretar com a obligacions d'actuació en una certa direcció i, per tant, com a una limitació de les possibilitats d'actuació. En altres paraules, això implica imposar restriccions sobre les transicions intra-escena i inter-escena de l'agent que adquireix els compromisos. Aquestes restriccions les modelitzarem amb regles d'inhibició. Una regla d'inhibició el que fa és impedir que una determinada transició es pugui realitzar. Això s'aconsegueix modificant la funció de transició de determinades escenes i prohibint certes transicions entre escenes. Per exemple, si volem eliminar una transició intraescena escriuriem $Inhibit(\rho(w, \varphi) = w')$ amb les particularitzacions necessàries de les variables d'agent a la fórmula φ.

Definition 4 *Donats un marc dialògic $DF = (Rols, L, ML, CL, Time)$ i una estructura performativa, $PS = (S, T, \alpha, \beta)$, una regla de comportament té la forma:*
If $\varphi_1 \wedge \ldots \wedge \varphi_n$ then $Inhibit(\psi) \in ML$
on $\varphi_i \in CL \cup L$ i ψ és de la forma $\rho(w, \varphi) = w'$ o bé $\psi \in T$

Finalment, podem agrupar els components d'una institució electrònica:

Definition 5 *Una Institució Electrònica és una tripleta $EI = (DF, PS, RB)$, on DF és un marc dialògic, PS és una estructura performativa i RB és un conjunt de regles de comportament.*

Actualment s'està desenvolupant una eina gràfica i un llenguatge per a l'especificació d'institucions mediades per agents. L'eina gràfica es troba descrita en detall a [1].

Només queda recordar que l'execució d'una institució consisteix en la generació d'instàncies d'escena (l'escena inicial en primer terme), l'assignació d'agents concrets a les variables d'agent d'una escena, i el moviment dels agents a través de les escenes respectant les restriccions de l'estructura performativa.

3 Aplicacions

3.1 Protocols mèdics

En la medicina moderna cada procés patològic porta associat un protocol. En aquest protocol s'especifica quines accions s'han de realitzar sobre els malalts en funció de quin és el seu estat i de quins han estat els resultats obtinguts a les proves realitzades. Cada vegada més s'intenta que l'acció dels metges vagi dirigida per aquests protocols i s'estableixen normes estrictes respecte del seu compliment.

Des del nostre punt de vista, la construcció d'agents que ajudin els metges a dirigir la seva actuació respectant els protocols establerts, o bé que permetin supervisar i monitoritzar la seva actuació, es un tema de recerca de gran interès. Forma part dels objectius d'un projecte CICYT anomenat SMASH en el que participa l'IIIA.

La línia de treball en curs per a modelitzar els protocols mèdics és la de considerar que un hospital és una institució dialògica, de la que són membres professionals de diferent tipus: metges, infermeres, auxiliars, caps de servei, directors mèdics, gerents, ... La interacció entre aquests professionals es realitza dins d'una estructura performativa rica, amb multitut d'escenes: visites als malalts, realització de proves, diàlegs elucidatoris entre professionals del mateix nivell, negociacions amb serveis farmacèutics, ... El que en medicina s'enten per "protocol" no és res més —ni menys— que l'estructura ordenada al llarg del temps de les diferents escenes (estructura performativa): primer s'ha de fer una ecografia, si el resultat es x, llavors fer una biopsia, ... i un conjunt de normes (regles de comportament) sobre què fer i què no fer en determinades circumstàncies. Actualment estem en procés de formalització del protocol de càncer de mama.

3.2 Comerç electrònic

Una de les àrees en les que hem estat treballant més intensament al llarg dels últims tres anys és la de comerç electrònic intermediat per agents. Es una de les àrees en la que els sistemes multi-agent tindran un impacte tecnològic més gran de manera immediata. Aquesta àrea és, de fet, la que ha inspirat les nocions bàsiques d'institució electrònica, ja que el comerç es realitza en la vida real, des de fa segles, intermediat.

L'aplicació on més hem treballat aquestes idees és la d'una casa de subhastes. En concret hem formalitzat la llotja de peix de Blanes [3]. Una casa de subhastes conté, de manera natural, totes les nocions d'una institució com les presentades en aquest treball. Determina l'ontologia (qualitat de peix, tipus de peix), identifica els participants (compradors, barques), fixa les escenes en les que els diferents participants dialoguen (admisió de compradors, de venedors, subhasta, pagament als pescadors, lliurament de mercaderies als compradors), els protocols associats (a la baixa en l'escena de subhasta), i les normes de comportament (penalitzacions per compres sense crèdit, expulsions de la llotja per mal comportament reiterat). El lector interessat pot revisar els treballs desenvolupats fins ara en aquesta àrea d'aplicació [5, 7, 6].

3.3 COMRIS

COMRIS és un projecte Europeu l'objectiu del qual és, bàsicament, establir formalismes que relacionin el món real amb un món virtual en que cada individu (real) és representat per un agent que realitza tasques mitjançant la interacció amb agents que representen altres individus. Un dels punts interessants del projecte és la interacció entre aquests dos móns, el real i el virtual. La interacció del món virtual cap al físic es farà amb un dispositiu físic portable que s'està desenvolupant actualment i que reb informació dels agents del món virtual via senyals de radiofreqüència. Aquest dispositiu sintetitza veu per ser sentida per l'individu que el porta. La interacció del món sensible al món virtual es farà mitjançant monitors, teclats y programes tradicionals.

El domini d'aplicació en el que s'està experimentant és el d'una conferència científica. Els assistents, en el moment de registrar-s'hi, són provistos del dispositiu físic i se'ls crea l'agent autònom. Està previst que l'agent ajudi l'individu, a partir de la informació sobre interessos que aquest li haurà proporcionat, a decidir a quines sessions assistir, a concertar cites amb persones amb interessos similars presents a la conferència, a organitzar reunions per a discutir propostes de projectes i tasques similars.

Les activitats que els agents realitzin en el món virtual poden ser vistes dins el marc d'una institució, novament. Bàsicament, el agents establiran diàlegs per a negociar dates per reunions, per a seleccionar quin agent té el perfil més adient per a ser un soci en un projecte europeu, etc. Aquests diàlegs poden estructurar-se molt còmodament mitjançant escenes, i es poden definir les regles de comportament bàsiques respecte dels compromissos que els agents hauran d'adquirir en les seves negociacions. Les primeres idees sobre l'aplicació d'institucions electròniques a una conferència es poden trobar a [4].

4 Discussió i treball futur

Les nocions presentades breument en aquest article són centrals per a la generalització de moltes aplicacions que actualment s'estan desenvolupant a laboratoris de recerca de manera ad-hoc, en particular en l'àrea de comerç electrònic. Actualment s'està en la fase inicial de desenvolupament del llenguatge d'especificació gràfic i textual. La generació automàtica d'institucions electròniques i d'agents que hi participarien són els objectius a llarg plaç d'aquesta recerca. Per exemple, ser capaços d'especificar una casa de subhastes amb els seus protocols i obtenir com a resultat tota la infraestructura de comunicacions i alguns dels agents participants: subhastador, admisor de venedors, supervisador i fins i tots agents prototípics de compradors i venedors, seria l'objectiu ambiciós d'aquest llenguatge d'especificació.

References

[1] T. Alsinet. JAFDIS, a java framework for dialogical institution specification. Technical report, Universitat de Lleida, 1998.

[2] P. Noriega and C. Sierra. Towards layered dialogical agents. In *Proceedings of the ECAI'96 Workshop Agents Theories, Architectures and Languages, ATAL'96*, number 1193 in LNAI, pages 157–171. Springer, 1996.

[3] Pablo Noriega. *Agent mediated auctions: The fishmarket metaphor.* PhD thesis, Universitat Autònoma de Barcelona, Barcelona, 1997.

[4] E. Plaza, P. Noriega, and C. Sierra. The conference centre as an agent-mediated institution. Technical report, IIIA-CSIC, 1998.

[5] J. A. Rodríguez, P. Noriega, C. Sierra, and J. Padget. A java-based electronic auction house. In *Second International Conference on The Practical Application of Intelligent Agents and Multi-Agent Technology: PAAM'97*, pages 207–224, London, UK, 1997.

[6] J. A. Ródriguez-Aguilar, F. J. Martín, P. Noriega, P. Garcia, and C. Sierra. Competitive scenarios forheterogeneus trading agents. In *Agents98*, pages 293–300, Minneapolis, MNN, 1998.

[7] Juan A. Ródriguez-Aguilar, Francisco J. Martín, Pablo Noriega, Pere Garcia, and Carles Sierra. Towards a test-bed for trading agents in electronic auction markets. *AICOMM*, page (in press).

[8] C. Sierra, N. R. Jennings, P. Noriega, and S. Parsons. A framework for argumentation-based negotiation. In M. P. Singh, A. Rao, and M. J. Wooldridge, editors, *Intelligent Agents IV*, number 1365 in LNAI, pages 177–192. Springer, 1997.