

3

INTERACCIÓN SOCIAL

El concepto de interacción social debe verse de manera análoga a las ideas introducidas previamente en esta primera parte del curso, i.e., así como el comportamiento autónomo y flexible es el fundamento de la agencia débil y la Intencionalidad es el de la agencia fuerte; la interacción social es la piedra angular de los SMA.

Primero es necesario acotar nuestro interés en el concepto de interacción. Aunque la idea general de que la dinámica de interacción entre componentes simples puede llevar a la emergencia de comportamiento inteligente [122], es generalmente aceptada, ésta suele concebirse a diferentes **niveles de abstracción**:

Niveles de interacción

- La relevancia de la interacción entre el agente y su medio ambiente ¹ ha sido discutida en la historia temprana de la IA por Simon [172] y más recientemente por la *nouvelle* IA reactiva [2]. En este nivel de abstracción, el concepto de interacción está más relacionado con la definición del comportamiento autónomo y flexible; así como la caracterización del medio ambiente en relación con los agentes ahí situados, i.e., agencia débil.
- El comportamiento de un agente puede verse como el resultado de las interacciones entre sus componentes. La Sociedad de la Mente, de Minsky [130], ejemplifica perfectamente este nivel; al igual que las redes de comportamiento de Maes [121]. En este nivel de abstracción, la interacción tiene que ver con la definición de arquitecturas de agente, independientemente de si nuestra noción de agencia es fuerte o débil.
- Las interacciones sociales, el nivel de nuestro interés en este capítulo, se dan, de acuerdo a Castelfranchi [32], cuando los participantes en la interacción son agentes en el sentido estricto del término –Sistemas Intencionales, posiblemente comunicativos.

La Figura 3.1 ilustra el origen de las **interacciones sociales** que pueden tener lugar a un SMA. La proyección en el medio ambiente representa la esfera de influencia y/o visibilidad de un agente según Jennings [109]. Esto significa que los agentes pueden, al menos potencialmente, influir en una parte del ambiente. El hecho de que las esferas de influencia presentan intersecciones, da lugar a interacciones entre los agentes, p. ej., dos agentes pueden intentar usar un mismo recurso. Las interacciones incluyen relaciones evidentes, como los canales de comunicación; pero también otras interacciones más sutiles como las relaciones de poder, p. ej., las existentes entre un jefe y su empleado.

Interacción social

¹ La misma idea es bien conocida en Etología, donde el comportamiento animal sólo puede entenderse en el contexto del ambiente en el que el animal habita. La idea ha sido elegantemente ilustrada por Braitenberg [20] en sus experimentos cibernéticos.

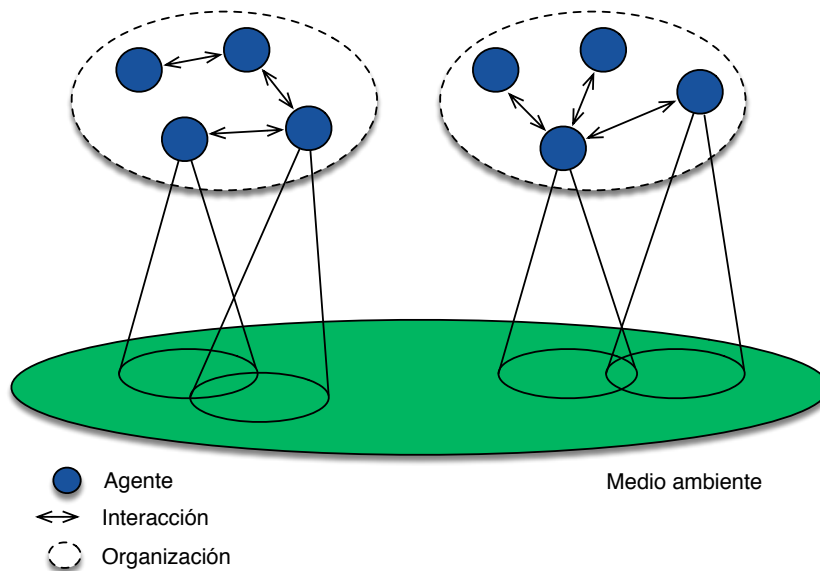


Figura 3.1: Agentes, interacciones y organizaciones en un SMA. Adaptado de Jennings [109]

El resto del capítulo hace una revisión detallada de los conceptos de situación de interacción y acción social, mismos que serán utilizados al abordar los temas propios del diseño e implementación de los SMA.

3.1 INTERACCIÓN

De acuerdo a Ferber [60], una **interacción** consiste en el establecimiento de una relación dinámica entre dos o más agentes mediante un conjunto de acciones recíprocas. La reciprocidad obedece a que las consecuencias de estas acciones, influyen el comportamiento futuro de los agentes participantes. Los participantes están de cierta forma en contacto, directo o indirecto, a lo largo de la interacción. De forma que esta noción de interacción asume:

Interacción

- La presencia de agentes capaces de actuar y/o comunicar.
- Situaciones susceptibles de servir de punto de encuentro entre los agentes: cooperación, negociación, regulación, etc., i.e., las intersecciones en la Figura 3.1.
- Elementos dinámicos que permiten las relaciones locales y/o temporales entre los agentes: comunicación, movilidad, etc.
- Autonomía, un juego que permita a los agentes establecer y abandonar este tipo de relaciones.

Morin [131] argumenta que, tal y como lo sugiere la Figura 3.1, las interacciones son la base constitutiva de una **organización**. Ahora bien, simultáneamente, estas interacciones suelen requerir de una organización predefinida para poderse llevar a cabo, p. ej., que los canales de comunicación adecuados estén establecidos –Las interacciones son fuente y producto de las organizaciones. De ahí su importancia para los SMA.

Organización

Una **situación de interacción** [60] es el conjunto de conductas que resulta de agrupar agentes que actúan para satisfacer sus objetivos, mientras que toman en consideración las restricciones producidas por recursos limitados y sus capacidades individuales. Situaciones donde los agentes son ayudados, dañados, favorecidos o amenazados sólo tienen sentido si tales agentes son sistemas provistos de alguna forma de **meta**. El término meta se usa aquí para denotar motivación: Deseos, intenciones, objetivos, propósitos, etc., independientemente de su representación como utilidades, valores, predicados, metas, etc.

Situación de Interacción

Metas y acción social

Dado que los agentes racionales tienen por definición **capacidades limitadas** y que muy seguramente enfrentan situaciones donde sus **recursos** son también limitados, la presencia de sus metas produce relaciones que constituyen las bases de la actividad social: situaciones de interacción. Una clasificación de las diferentes situaciones de interacción posibles es ofrecida por Ferber [60], con base en:

Capacidades y Recursos limitados

- **Las metas de cada agente.** Las metas de los agentes pueden ser compatibles o incompatibles. La meta ϕ de un agente ag_1 es incompatible con la meta ψ de un agente ag_2 , si $\phi \rightarrow \neg\psi$, i.e., si el agente ag_1 satisface su meta, ag_2 no lo hará. Esta situación introduce una primera clasificación: Si varios agentes tienen metas compatibles, se encuentran en una situación cooperativa o una de indiferencia; si tienen metas incompatibles, se encuentran en una situación antagónica.
- **Recursos.** Todo agente necesita recursos para ejecutar sus acciones. La cantidad disponible de estos recursos, normalmente es limitada. Esto es fuente de conflictos: Algunos agentes necesitarán el mismo recurso, al mismo tiempo, en el mismo lugar. Se utilizan mecanismos de coordinación para evitar la ocurrencia de conflictos. Se utilizan mecanismos de negociación para resolver aquellos conflictos que lleguen a presentarse.
- **Las habilidades de cada agente.** Dadas las habilidades o capacidades de un agente, existen dos casos a considerar. Primero, el agente puede satisfacer sus metas sólo; Segundo, el agente necesita de otros agente para lograr satisfacer sus metas. En el segundo caso, el sistema que resulta de las interacciones entre agentes, tiene nuevas propiedades que se consideran funcionalidad emergente.

Bajo **situaciones cooperativas**, en donde los agentes tienen metas compatibles, las siguientes situaciones de interacción son posibles:

Situaciones cooperativas

- **Independencia.** Los agentes tienen suficientes recursos y habilidades para satisfacer sus metas. El nivel SMA consiste en la ejecución de las acciones individuales de los agentes, de manera independiente. De hecho, propiamente hablando, no hay una interacción entre ellos. Como ejemplo consideren un grupo de ingenieros trabajando en el mismo proyecto.
- **Coordinación simple.** Los agentes tienen suficientes recursos, pero sus habilidades son insuficientes. Es necesario, por lo tanto, sumar sus habilidades para lograr satisfacer sus metas. Este es el caso clásico de los

problemas que involucran varios especialistas, donde la interacción es expresada a través de la asignación de tareas y el conocimiento compartido. No se necesitan formas suplementarias de coordinación.

- **Congestión.** Los agentes tienen suficientes habilidades para satisfacer sus metas, pero los recursos disponibles son insuficientes. En este caso, los agentes se obstruyen al tratar de cumplir con sus metas. El ejemplo más claro en este caso, es manejar en una autopista congestionada; o el financiamiento en tiempos de crisis.
- **Colaboración coordinada.** Tanto los recursos, como las habilidades de los agentes, son insuficientes para lograr la satisfacción de las metas. Este escenario constituye la situación cooperativa más compleja, donde se requiere considerar coordinación en el uso de recursos y el problema de asignación de tareas. Como ejemplo puede incluirse todo proceso industrial.

En **situaciones antagónicas** pueden presentarse los siguientes casos de interacción:

*Situaciones
antagónicas*

- **Competencia individual pura.** Los agentes tienen suficientes recursos y habilidades, pero sus metas son incompatibles. Tendrán que negociar o luchar para lograr sus metas. La palabra “pura” se refiere al hecho de que el acceso a los recursos es igual para todos los agentes. En este caso, el nivel SMA puede verse como “el mejor gana”, sin interacción efectiva.
- **Competencia colectiva pura.** Los agentes tienen suficientes recursos, pero sus habilidades son insuficientes. Tendrán que formar grupos o coaliciones para lograr cumplir sus metas. Los grupos que emergen están sujetos a una competencia similar a la competencia individual pura.
- **Conflicto individual por recursos.** Los agentes tienen suficientes habilidades, pero los recursos son insuficientes. Puesto que los recursos no pueden ser compartidos, cada agente querrá adquirirlos para satisfacer sus metas.
- **Conflicto colectivo por recursos.** En este caso, los agentes no tienen las habilidades suficientes, ni los recursos necesarios para lograr satisfacer sus metas. Esta situación puede verse como una combinación de la competencia colectiva y el conflicto individual por recursos. Los grupos emergentes pelearán por monopolizar los recursos. La guerra y la competencia industrial son ejemplos de tal situación.

El Cuadro 3.1 resume esta clasificación. En lo que sigue estudiaremos una aproximación particular a las acciones sociales que componen las situaciones de interacción.

Situaciones de Interacción Cooperativas			
Interacción	Metas	Habilidades	Recursos
Independencia	compatibles	suficientes	suficientes
Coordinación simple	compatibles	insuficientes	suficientes
Congestión	compatibles	suficientes	insuficientes
Colaboración coordinada	compatibles	insuficientes	insuficientes
Situaciones de Interacción Antagónicas			
Interacción	Metas	Habilidades	Recursos
Competencia individual pura	incompatibles	suficientes	suficientes
Competencia colectiva pura	incompatibles	insuficientes	suficientes
Conflicto individual por recursos	incompatibles	suficientes	insuficientes
Conflicto colectivo por recursos	incompatibles	insuficientes	insuficientes

Cuadro 3.1: Situaciones de interacción clasificadas con base en metas, habilidades y recursos disponibles de los agentes participantes.

3.2 ACCIÓN SOCIAL

Castelfranchi [32] define una **acción social** como la acción ejecutada por un agente al asumir, en el marco de una situación de interacción, que los otros participantes en la interacción también son agentes inteligentes ².

Acción social

Observen que esta definición no hace énfasis en lo **colectivo**, de hecho, define la acción **individual** de un agente en el contexto de una situación de interacción. Esto es, una acción colectiva no es necesariamente social. Por ello es necesaria una definición de acción social individual. Por ejemplo, la acción colectiva de ir a trabajar todos los días, difícilmente puede verse como una acción social, a diferencia de lo que sucede con las acciones individuales de los participantes de una huelga general.

Colectivo vs Social

El propósito y la **Intencionalidad** de un agente determinan si una acción es social, o no lo es. La noción de acción social no puede estar basada exclusivamente en el comportamiento, esto es, en una descripción externa. Es necesario tomar en cuenta los estados mentales de los agentes, incluyendo su representación de otros agentes. No es de extrañar pues, que Castelfranchi y su teoría Intencional de acción social sean nuestra elección para este tema. Si nuestro enfoque fuese más cercano a la *nouvelle IA*, basada en el comportamiento, el referente a acción social sería más cercano al conductivismo. De hecho, Castelfranchi considera que lo cognitivo y lo emergente en estas propuestas no son nociones incompatibles, ni mucho menos enfoques necesariamente alternativos en el estudio de la inteligencia y la cooperación. Lo cognitivo debe estudiarse como una propiedad emergente (de lo sub-simbólico a lo simbólico, de lo objetivo a lo subjetivo, de lo implícito a lo explícito). Los fenómenos emergentes como la inteligencia de enjambre, etc., no debería modelarse y estudiarse exclusivamente desde la perspectiva de agentes sub-cognitivos.

Intencionalidad vs Conductismo

² Castelfranchi considera explícitamente que los agentes exhiben un comportamiento dirigido por metas, o con propósito (*purposive behavior*) en términos de Rosenblueth y Wiener [159], en cuyo trabajo las metas comparten algunas propiedades de las intenciones à la Bratman: guías y controladores de la acción, factor determinante de la función de selección de acción.

Sobre el propósito de los agentes, es importante mencionar que una acción dirigida hacia otro agente, no es necesariamente una acción social. También es cierto que una acción que no está dirigida hacia ningún agente, puede ser una acción social. Esto se debe a que son las **metas**, y no las diferencias de comportamiento, las que determinan si una acción es social o no. Por ejemplo, un agente que cierra una puerta para evitar que otros agentes entren en cierta área, lleva a cabo una acción social. La misma acción llevada a cabo por que el agente va a pintar la puerta, no es una acción social.

Propósito

Así como una acción colectiva no es necesariamente social, tampoco puede reducirse ésta a la **comunicación**. Los agentes no son sociales porque comunican, sino que comunican por ser sociales. La comunicación obedece a que los agentes interfieren, dependen e influyen a otros agentes en su ambiente. De estas interacciones, la inclusión de la interferencia debería llamarnos la atención: las acciones sociales, comunicación incluida, no son necesariamente una actividad cooperativa.

Comunicación y acción social

En lo que sigue presentaremos la ontología básica de acciones sociales, propuesta por Castelfranchi. La presentación está organizada como si fuésemos a construir incrementalmente el concepto de lo social.

3.2.1 Interferencia y Dependencia

Las acciones sociales pueden clasificarse usando el concepto de **interferencia**, es decir, situaciones donde el efecto de una acción de un agente es relevante para las metas de otro agente. Las interferencias pueden clasificarse en negativas y positivas. Las negativas se dan cuando una acción amenaza o compromete el cumplimiento de las metas de otros agentes. Las positivas se dan cuando la acción favorece el cumplimiento de las metas de otros agentes. Observen que esto es similar a la clasificación inicial de las situaciones de interacción en la sección anterior, pero en ese caso la relación era entre las metas de los agentes, mientras que aquí es entre las acciones y las metas de los agentes.

Interferencia

La **dependencia** se define como un caso de interferencia fuerte, donde un agente no puede satisfacer sus metas porque necesita una acción o recurso controlados por otro agente. Observen que la dependencia se define en términos de habilidades y recursos. Es evidente que las situaciones de interacción cooperativas pueden verse como interferencias positivas, mientras que las situaciones antagónicas pueden verse como interferencias negativas.

Dependencia

Ahora bien, desde una perspectiva egocéntrica, un agente ag_1 tiene dos alternativas al confrontar situaciones de interferencia y dependencia con un agente ag_2 :

1. ag_1 puede adaptar su comportamiento al del agente ag_2 , con el fin de explotar las acciones de ag_2 ; o para evitar interferencias negativas con este agente.
2. ag_1 puede tratar de cambiar el comportamiento de ag_2 induciéndolo a actuar de la manera deseada; o a que abandone su comportamiento pernicioso.

Estas alternativas definen las acciones sociales elementales, como la mera coordinación, la influencia y la delegación, que se definen en el resto de esta sección. El Cuadro 3.2 resume estas alternativas. La columna Adaptarse, representa la mera coordinación, ya sea positiva o negativa; la columna Inducir, representa la delegación. En ambos casos tenemos posiblemente acciones sociales del agente ag_1 , pero de muy distinta naturaleza. Además, como veremos, las acciones de ag_1 son sociales solo bajo ciertas condiciones específicas.

Interferencia	Adaptarse	Adaptar al otro
Negativa	ag_1 modifica su plan para evitar interferencia.	ag_1 induce a ag_2 a evitar la interefencia.
Positiva	ag_1 modifica su plan incluyendo una acción de ag_2 .	ag_1 induce a ag_2 a perseguir la meta de ag_1 .

Cuadro 3.2: Acciones sociales con base en el tipo de interferencia.

Toda acción puede verse como una interacción, debido a que el ambiente no es una entidad pasiva. La acción de un agente afecta el ambiente y la retroalimentación en el ciclo de razonamiento de un agente viene del ambiente. De hecho, hay una forma de delegación en el ambiente, en sus dinámica y su causalidad. Así que las acciones son interacciones entre el agente y su ambiente, aunque claro, éstas **no son** interacciones sociales –el ambiente no es un agente. Por ejemplo, es racional explotar al sol para conseguir energía, pero es irracional pedirle ayuda para ello.

Acciones no sociales

Una **acción social** es una acción que contiene con otra entidad como si ésta fuese un agente. Las interacciones con otros seres vivos comienzan a ser sociales. Si bien el sol no colabora con nosotros, una planta sí al darnos frutas. La agricultura es una forma de colaboración entre el ser humano y las plantas. Para un agente Intencional una acción social es aquella que contiene con otro agente con base en la postura Intencional [48]: con base en una representación de las creencias, deseos e intenciones del otro. De hecho, haremos una distinción entre **acción social débil**, aquella que se basa solo en creencias, aunque sean creencias acerca de otro agente; y **acción social fuerte**, aquella que se basa o está dirigida por metas sociales. De ahí la relevancia de la Intencionalidad en la acción social.

Acción social

Acción social débil

Acción social fuerte

Las creencias acerca del estado mental de otro agente no son sólo resultado de la comunicación, o del razonamiento basado en estereotipos; sino de la **interpretación** del comportamiento del otro agente. El comportamiento del otro es un signo de su propio estado mental, p. ej., si un agente entrega una solicitud de ingreso a nuestra maestría, puedo interpretar desde la postura Intencional que intenta cursarla. Este entendimiento es, antes que la comunicación explícita, la base de la coordinación y colaboración recíprocas. De ahí la relevancia de la postura Intencional en la acción social.

Postura Intencional

3.2.2 La coordinación pura

Restringiremos el concepto de coordinación a situaciones donde un agente ag_1 coordina su comportamiento con el comportamiento observado o predicho de otros agentes, ignorando la posibilidad de cambiar el comportamiento de esos agentes. Esto es la **coordinación pura** sin influencia, ni comunicación. Nuestra intuición de coordinación suele incluir formas como la negociación, la cooperación, la resolución de conflictos, etc. [7, 106]. Aquí nos referiremos exclusivamente a la coordinación pura, que puede tomar diferentes formas:

- **Coordinación reactiva vs anticipada.** La manera en que un agente detecta las interferencias define dos tipos de coordinación: En la **reactiva**, el agente usa directamente su percepción para identificar obstáculos, oportunidades y la reacción adecuada a ellos; En la **anticipada** el agente se basa en el aprendizaje o la inferencia, requiriendo de un modelo del medio ambiente. La coordinación anticipada reduce riesgos al ganar tiempo para adaptar el comportamiento del agente a la situación esperada. Los agentes reactivos pueden exhibir cierta forma de coordinación anticipada bajo la forma de **reflejos condicionados**. Castelfranchi argumenta que ningún agente puede realmente planear su comportamiento sin un mecanismo de coordinación anticipada.

Coordinación pura

Coordinación reactiva

Coordinación anticipada

Reflejos condicionados
- **Coordinación positiva vs negativa.** Dependiendo del tipo de interferencia presente en la situación de interacción, dos tipos de coordinación se presentan: La coordinación **negativa**, o evitación, se debe a interferencias negativas y se orienta a evitar un daño u obstáculo; En la coordinación **positiva** o explotación, los agentes llevan a cabo un cambio en los planes propios, asignando parte de ellos a otro agente para sacar provecho de una situación social favorable, lo que se conoce como **delegación**.

Coordinación negativa

Coordinación positiva

Delegación
- **Coordinación egoísta vs colaboración.** En la coordinación **egoísta** o ego-céntrica, un agente ag_1 trata de satisfacer sus propias metas ante la presencia de otros agentes. En la **colaboración**, el agente ag_1 adapta su comportamiento para favorecer a otros agentes. Observen que la colaboración constituye una acción social fuerte, debido a que no esta basada únicamente en las creencias sobre el estado mental de otros agentes, sino que es guiada por una meta social: que el otro agente logre la meta de ag_1 o colabore a ello. Esto implica una forma de ayuda que puede ser activa o pasiva.

Coordinación egoísta

Colaboración
- **Coordinación unilateral, bilateral y mutua.** Dependiendo del conocimiento sobre la presencia de otros agentes en la coordinación, es posible considerar tres casos: Dados dos agentes, ag_1 y ag_2 involucrados en una situación de coordinación, hablamos de coordinación **unilateral** si sólo ag_1 está coordinando su comportamiento con el del agente ag_2 . Es posible que el agente ag_2 esté haciendo lo mismo y entonces se habla de una coordinación **bilateral**. Si ambos agentes conocen que están coordinados y tratan de llegar a algún acuerdo se habla de coordinación **mutua**. En estos casos se requiere alguna forma de colaboración.

Coordinación unilateral

Coordinación bilateral

Coordinación mutua

3.2.3 Delegación

En la **delegación**, un agente ag_1 necesita de, o le gustaría incluir, una acción del conjunto de habilidades de ag_2 en su propio plan de acciones. La delegación se debe a la **dependencia**: El agente ag_1 para poder satisfacer una meta que no puede cumplir por sí mismo, delega una acción al agente ag_2 , que puede y desea ejecutar esa acción. Si el agente ag_1 tiene conocimiento de las acciones del agente ag_2 , estará delegando acciones que son útiles para él. De alguna manera, el agente ag_1 **confía** en el agente ag_2 , lo que significa que el planea lograr una meta γ a través de la acción del agente ag_2 . Las siguientes condiciones caracterizan la **confianza** del ag_1 en el ag_2 :

Delegación

Dependencia

Confianza

- El agente ag_1 cree que el agente ag_2 puede y desea hacer una acción dada;
- El agente ag_1 tiene la meta de que el agente ag_2 haga esa acción; y
- El agente ag_1 se abstiene de ejecutar la acción delegada a ag_2 , coordinando su comportamiento con el del agente ag_2 .

Diferentes **tipos de delegación** son posibles, incluyendo formas sociales y no sociales de delegación:

Tipos de delegación

- **Delegación débil unilateral.** En este caso, no hay conocimiento bilateral de la delegación, ni acuerdo alguno. El agente ag_2 no conoce que el agente ag_1 está explotando su acción y el agente ag_1 sólo explota el comportamiento autónomo de ag_2 . Por ejemplo, un cazador dispara su flecha asumiendo que su presa se seguirá moviendo en la misma dirección. La presa colabora débilmente con el cazador; el profesor y su sistema de evaluación.
- **Delegación por inducción.** Esta es una forma más fuerte de delegación donde el agente ag_1 provoca o induce el comportamiento del agente ag_2 para explotarlo. La base de la inducción depende de la naturaleza del agente inducido y puede incluir estímulos, creencias, influencia, etc. Por ejemplo, el pescador y su pez.
- **Delegación por aceptación.** En este caso, también conocido como **delegación fuerte**, el agente ag_2 conoce la intención del agente ag_1 de explotar su acción. Normalmente se basa en la aceptación por parte del agente ag_2 de adoptar la meta del agente ag_1 (adopción social de metas), posiblemente tras una negociación concluida en un acuerdo o compromiso social.

Para los agentes cognitivos que utilizan metas y planes para satisfacerlas, la delegación puede especificarse en diferentes **niveles** de abstracción. Los más importantes son:

Niveles de delegación

- **Delegación ejecutiva pura vs abierta.** En la delegación abierta, la tarea delegada se especifica mínimamente. En la delegación ejecutiva pura, la tarea se especifica completamente. Obviamente, tal especificación

puede darse a un nivel intermedio. Observen que la delegación ejecutiva es diferente a la programación directa, inaceptable en términos de la autonomía de los agentes.

- **Delegación de tareas en el dominio vs tareas de planeación.** Un agente puede delegar tareas que se relacionan con el dominio del problema, así como meta-acciones relacionadas con tareas de planeación.

La delegación abierta asegura la flexibilidad de los planes distribuidos en un SMA, dado que una especificación completa de tareas normalmente no se puede ejecutar o presenta inconvenientes. El carácter distribuido de estos planes se derive de la delegación abierta. La combinación de la delegación parcial y abierta crea la posibilidad de que los agentes cooperen en planes que no comparten y nadie conoce enteramente: un **plan distribuido**. ag_1 puede delegar a ag_2 un plan entero, o bien una parte de un plan en lo que se conoce como **delegación parcial**. Observen que aunque el plan no es conocido por todos los agentes, todas las partes de éste son conocidas por al menos un agente. Esta es la base de la **cooperación orquestada**, del estilo jefe decidiendo un plan general. Sin embargo, esa base es insuficiente para la emergencia de cooperación funcional no consciente entre agentes planificadores.

Plan distribuido

Delegación parcial

Cooperación orquestada

3.2.4 Acción social fuerte

En la delegación, el agente ag_1 tiene la meta de que ag_2 lleve a cabo una acción dada que necesita incluir en su plan. Si ag_2 es un agente cognitivo, entonces ag_1 también tiene la meta de que ag_2 intente llevar a cabo esa acción. Esto es lo que se conoce como **delegación cognitiva**, el acto de delegar en un agente Intencional. Esta meta de ag_1 es una razón para influenciar a ag_2 , pero no necesariamente ag_1 tiene que influenciar a ag_2 . Puede ser el caso que ag_1 no tenga que hacer nada porque ag_2 independientemente intenta llevar a cabo la acción que ag_1 necesita.

Delegación cognitiva

Una acción social fuerte está caracterizada por metas sociales. Una **meta social** es una meta que está dirigida hacia otro agente. Por ejemplo, Eva quiere que Adán crea que no se comió la manzana. Eva quiere que Adán le traiga más manzanas. Las combinaciones de intenciones personales y creencias sobre las intenciones de los demás son insuficientes para el entendimiento de las interacciones sociales, colaboración y organización; para ello es necesario el concepto de meta social.

Meta social

La vida social de los agentes cognitivos está más allá de la coordinación pura: ¿Cómo inducir a un agente a que crea algo o haga algo por nosotros? ¿Cómo inducirle a que no haga algo? Normalmente, pero no necesariamente, comunicando. El problema es que cuando comunicamos solo podemos informar a otro agente acerca de nuestras metas y nuestras creencias acerca de sus acciones, pero ¿Porqué debería el otro agente preocuparse por nuestras metas y expectativas?

Hemos asumido que el otro agente es o bien **benevolente**, o bien un esclavo obediente, pero esto no es necesariamente el caso. Para poder inducirle a hacer o no hacer algo, necesitamos **poder** sobre él. Su benevolencia ha-

Benevolencia

Poder

cía nosotros (simpatía, etc.) es sólo una de las posibles bases del poder que nos permite influenciar su comportamiento. Sin embargo, la fuente de poder más importante en este caso, es el hecho de que nuestras acciones interfieren potencialmente con sus metas: dependencia y recompensa.

La norma general para influenciar a un agente cognitivo no se basa en la ingeniería de incentivos, sino en la modificación de las creencias que dan soporte a sus metas e intenciones. Se trata de proveer razones para actuar de una manera y no de otra. Las creencias sobre incentivos, por ejemplo, las **funciones de utilidad**, solo son un caso particular de la norma general.

Preferencias

3.2.5 Adopción social de metas

En la **adopción de metas** la mente del agente ag_2 está cambiando. El agente adquiere una nueva meta, o al menos una nueva razón para una meta que tenía previamente. La razón para esta nueva meta es que el agente ag_1 quiere satisfacerla. El agente ag_2 sabe esto y decide hacer, o dejar que el otro agente haga, que la meta se cumpla. El agente ag_2 adquiere la misma meta que el agente ag_1 porque sabe que se trata de una meta del agente ag_1 .

Adopción de metas

La adopción social de metas no debe confundirse con la **imitación** que es una acción social unilateral similar. Cuando ag_2 imita a ag_1 no intenta que ag_1 vea cumplida su meta γ , solo quiere resolver su meta γ' que es igual o similar a la de ag_1 ; en cambio, cuando ag_2 adopta la meta γ de ag_1 , ag_2 quiere que ag_1 satisfaga γ al menos mientras ag_2 cree que ag_1 desea satisfacerla.

Imitación

La adhesión de metas, o **conformidad**, es la forma más fuerte de adopción social de metas. Ocurre cuando una adopción social de metas responde a una petición. Es la forma opuesta a la adopción espontánea de metas. No sólo el agente ag_1 tiene la meta γ y el agente ag_2 adopta esta meta, sino que ag_1 tiene como meta que ag_2 haga algo para el cumplimiento de γ , y hacer saber a ag_2 de sus expectativas. Para satisfacer al agente ag_1 , el agente ag_2 no sólo debe lograr γ , además debe hacer saber a ag_1 que ejecutará la acción delegada y que producirá γ . El acuerdo está basado en la adhesión. La delegación fuerte es un requisito para la adhesión.

Conformidad

La adopción de metas puede estar motivada por las distintas formas de benevolencia: piedad, altruismo, amor, amistad, etc., pero también puede ser **instrumental**. Por ejemplo, alimentar pollos es un medio para eventualmente hacer caldo. La adopción de metas instrumental sucede también en el intercambio social. Otra forma de adopción de metas motivada es la **adopción cooperativa** de metas donde el agente ag_2 adopta la meta de ag_1 porque se interesa también en algunos de los resultados intentados por ag_1 . La distinción entre estas formas de adopción de metas es importante pues su base emotiva diferenciada permite importantes predicciones sobre la conducta cooperativa de ag_2 .

Adopción de metas instrumental

Adopción cooperativa de metas

Como sucede con la delegación, diferentes dimensiones pueden ser caracterizadas en el caso de la adopción social de metas. Dados dos agentes, ag_1 que adopta una meta y ag_2 que la delega, se pueden considerar los siguientes **niveles** de adopción de metas:

Niveles de adopción de metas

- **Ayuda literal.** El agente ag_1 adopta exactamente lo delegado por el agente ag_2 .

- **Sobre-ayuda.** El agente ag_1 va más allá de lo delegado por el agente ag_2 , sin cambiar el plan del agente ag_2 .
- **Ayuda crítica.** El agente ag_1 satisface los resultados relevantes de la acción o el plan delegado, pero lo modifica.
- **Ayuda sobre-crítica.** El agente ag_1 lleva acabo una sobre-ayuda, modificando el plan del agente ag_2 .
- **Ayuda hiper-crítica.** El agente ag_1 adopta metas o intereses del agente ag_2 , que el mismo agente ag_2 no ha considerado. Al hacer esto, el agente ag_1 no lleva acabo la acción o plan solicitado, ni satisface los requerimientos que le fueron delegados.

Un agente realmente útil debería tomar en cuenta nuestras metas e intereses, e ir más allá de nuestras delegaciones y peticiones. Castelfranchi asume que sólo los agentes cognitivos pueden ayudar de manera no accidental, o al menos conscientemente, más allá de la delegación al reconocer nuestras necesidades actuales y contextuales.

3.2.6 Metas sociales y compromiso social

En esta teoría de acción social y agencia, las **metas sociales** que están en la mente de los miembros del grupo son el verdadero pegamento de la actividad conjunta. Ninguna actividad de grupo, ni plan conjunto, ni verdadera colaboración puede establecerse sin:

Metas sociales

- La meta del agente ag_1 (miembro del grupo) acerca de la intención del agente ag_2 de hacer una tarea o acción dada α (delegación);
- La creencia del agente ag_1 de que ag_2 es capaz y tendrá la oportunidad de hacer α .
- El compromiso social del agente ag_2 con el agente ag_1 para α , el cual es una forma de adopción de metas o una mejor adhesión.

El **compromiso social** resulta de mezclar una delegación fuerte y su correspondiente adopción fuerte. El compromiso social recíproco constituye la estructura más importante de cualquier grupo u organización.

Compromiso social

Hay una forma pre-social del compromiso, llamada compromiso interno o individual, que se da entre un agente y una acción cuando: El agente ha decidido hacer algo, el agente está determinado a ejecutar una acción dada a un tiempo dado, y la meta es persistente (recuerden las formas de compromiso en los agentes BDI).

Un compromiso social no es un compromiso individual compartido por varios agentes, es un concepto relacional [31]: el compromiso de un agente con otro. De hecho es una relación de cuatro argumentos: ag_1 es el agente comprometido; a es la acción a la cual ag_1 se ha comprometido; ag_2 es el agente con quien ag_1 se ha comprometido; y ag_3 es un posible testigo o garante del compromiso de ag_1 .

El compromiso colectivo o de grupo es un compromiso interno de un agente colectivo o grupo hacia una meta colectiva. Requiere del compromiso social de sus miembros hacia otros miembros del grupo.

El compromiso social combina la delegación por aceptación y la adopción por aceptación, pero además cuando el agente ag_1 está socialmente comprometido con otro agente ag_2 , el agente ag_2 puede: controlar si el agente ag_1 hace lo prometido; pedirle que lo haga; quejarse o protestar ante ag_1 , si no hace lo prometido; y en algunos casos hacer efectivas sus compensaciones ante pérdida, regalías, etc. Por lo tanto, el compromiso social crea **derechos y responsabilidades** entre los agentes ag_1 y ag_2 .

Derechos y responsabilidades

3.2.7 Estructura social y organización

El compromiso social no es la única estructura importante que restringe la actividad social. La coordinación en un grupo no está únicamente garantizada por un estado mental compartido (intenciones conjuntas, planes acordados, creencias compartidas), benevolencia recíproca o comunicación. Existen diferentes **estructuras sociales** en cualquier SMA:

Estructuras sociales

1. La interdependencia y la estructura de poder;
2. La estructura de conocidos, que emerge de la unión de todos los conocidos de cada agente;
3. La estructura de comunicación (la red global de los posibles canales de comunicación directos o indirectos);
4. La estructura de compromiso, que emerge de todas las relaciones de delegación-adopción y la formación de coaliciones;
5. La estructura determinada y preestablecida de reglas y normas sobre acción e interacción.

Estas estructuras no sólo afectan la posibilidad de éxito de los agentes, sino que, si son conocidas, restringen sus decisiones, metas y planes. Algunas de esas estructuras son construidas deliberadamente por los agentes, mientras que otras emergen de manera objetiva.

La estructura de interferencia e interdependencia entre una población de agentes, es emergente y objetiva. Además es independiente del conocimiento de los agentes y sus decisiones. Esta estructura restringe las acciones de los agentes al determinar su éxito o eficacia. Dado un grupo de agentes en un medio ambiente común, y dadas sus metas y sus diferentes y limitadas habilidades y recursos, estos de hecho dependen entre sí. Es así como la estructura de dependencia emerge. Esta red de dependencia determina y predice alianzas y coaliciones, competición, cooperación, intercambio, estructura funcional en organizaciones, comunicación racional o efectiva, y poder de negociación.

Esta estructura pre-cognitiva, puede emerger cognitivamente en la medida en que las restricciones que impone se vuelven conocidas para los agentes, ya sea por su comprensión o por aprendizaje ciego. La estructura objetiva de interdependencias retro-alimenta a los agentes, cambiando así su estado

mental. La emergencia de esta estructura auto-organizada, no decidida y no contractual, es necesaria para derivar metas sociales.

3.3 LECTURAS Y EJERCICIOS SUGERIDOS

Gasser [68] introduce los conceptos de conocimiento y acción sociales, como fundamento de la IA distribuida y abierta. Algunas de las ideas aquí discutidas, están también presentes en el trabajo de Sichman y col. [171] sobre mecanismos sociales bajo redes de dependencia. En el mismo sentido, un estudio más reciente de la implicación de la estructura social en agentes software es presentada por Ossowski [141]. Una discusión más detallada de la acción social puede encontrarse en el libro de Conte y Castelfranchi [42]. El concepto de confianza es analizado en detalle por Castelfranchi y Falcone [29].

Ejercicios

Ejercicio 3.1. *Ejemplifique un caso de situación cooperativa y uno de situación antagónica. Justifique sus ejemplos.*

Ejercicio 3.2. *Ejemplifique la diferencia entre acción social débil y fuerte.*

Ejercicio 3.3. *Ejemplifique un caso de delegación.*

Ejercicio 3.4. *Ejemplifique un caso de imitación.*

Ejercicio 3.5. *Una posible forma de aprendizaje intencional [84, 87], consiste en que un agente adopta la meta de aprender un concepto, mientras que otros agentes interesados en el mismo concepto adoptan la meta de recolectar ejemplos de aprendizaje sobre el concepto en cuestión a cambio de lo aprendido. ¿Qué clase de acción social es esta forma de aprendizaje?*