

Modelos simbólicos de IA Social

Sistemas Inteligentes Distribuidos

Sergio Alvarez

Javier Vázquez

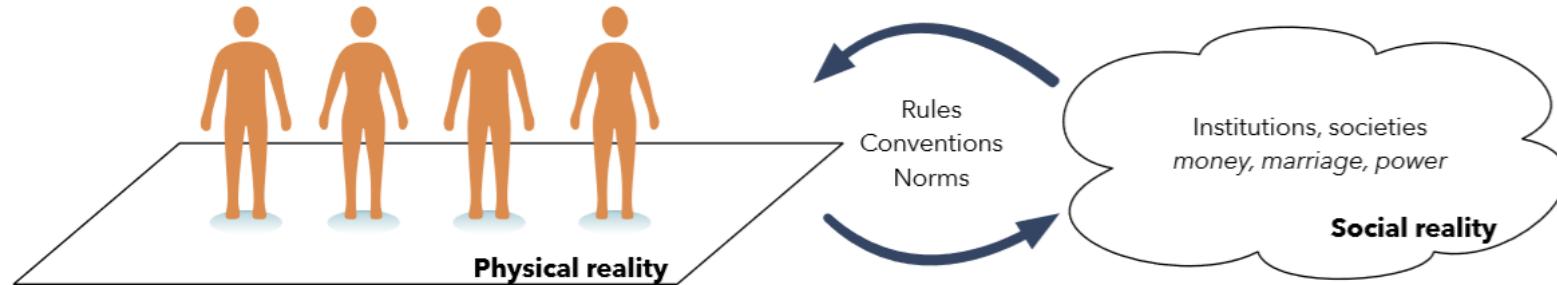
Bibliografía

- N.V. Findler and R. Malyankar. “*Social structures and the problem of coordination in intelligent agent societies*”. 2000.
- Virginia Dignum, “*A Model for Organizational Interaction: Based on Agents, Founded in Logic*”. PhD Thesis, Universiteit Utrecht, 2004
- J. Vázquez Salceda. “*The Role of Norms and Electronic Institutions in Multiagent Systems*”, Birkhauser-Verlag, 2004
- J. Vázquez-Salceda, V. Dignum, F. Dignum. “*Organizing Multiagent Systems*”. *Journal of Autonomous Agents and Multi-Agent Systems*, vol. 11 issue 3, pp. 307-360, 2005.
- S. Alvarez-Napagao. “*Bringing Social Reality to Multiagent and Service Architectures*”. PhD Thesis, Universitat Politecnica de Catalunya-BarcelonaTECH, 2016.

Introducción

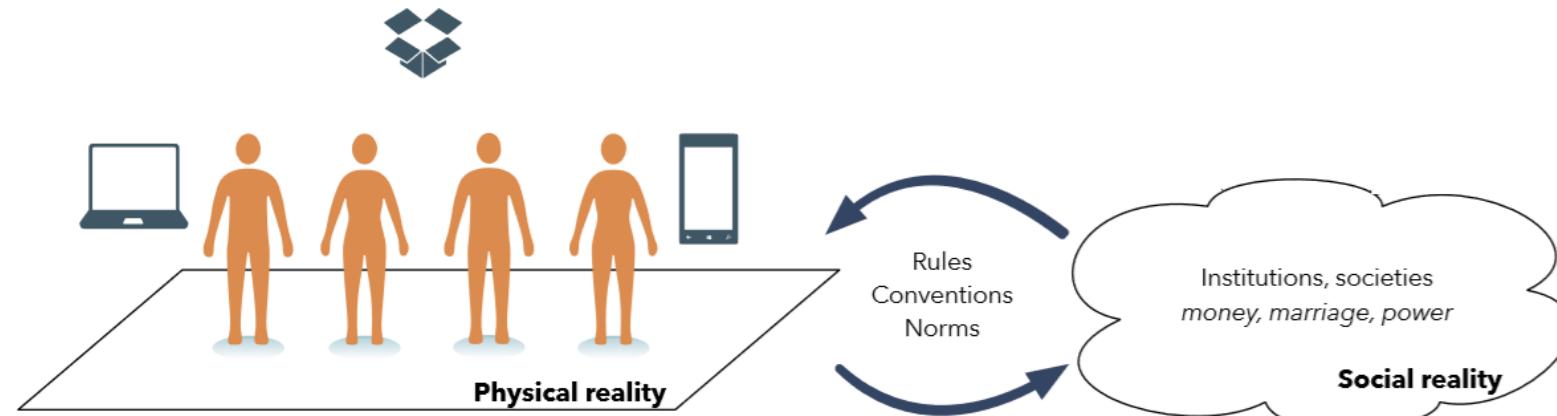
Modelos simbólicos de IA Social

Impacto en la sociedad: La IA en los Sistemas Sociotécnicos



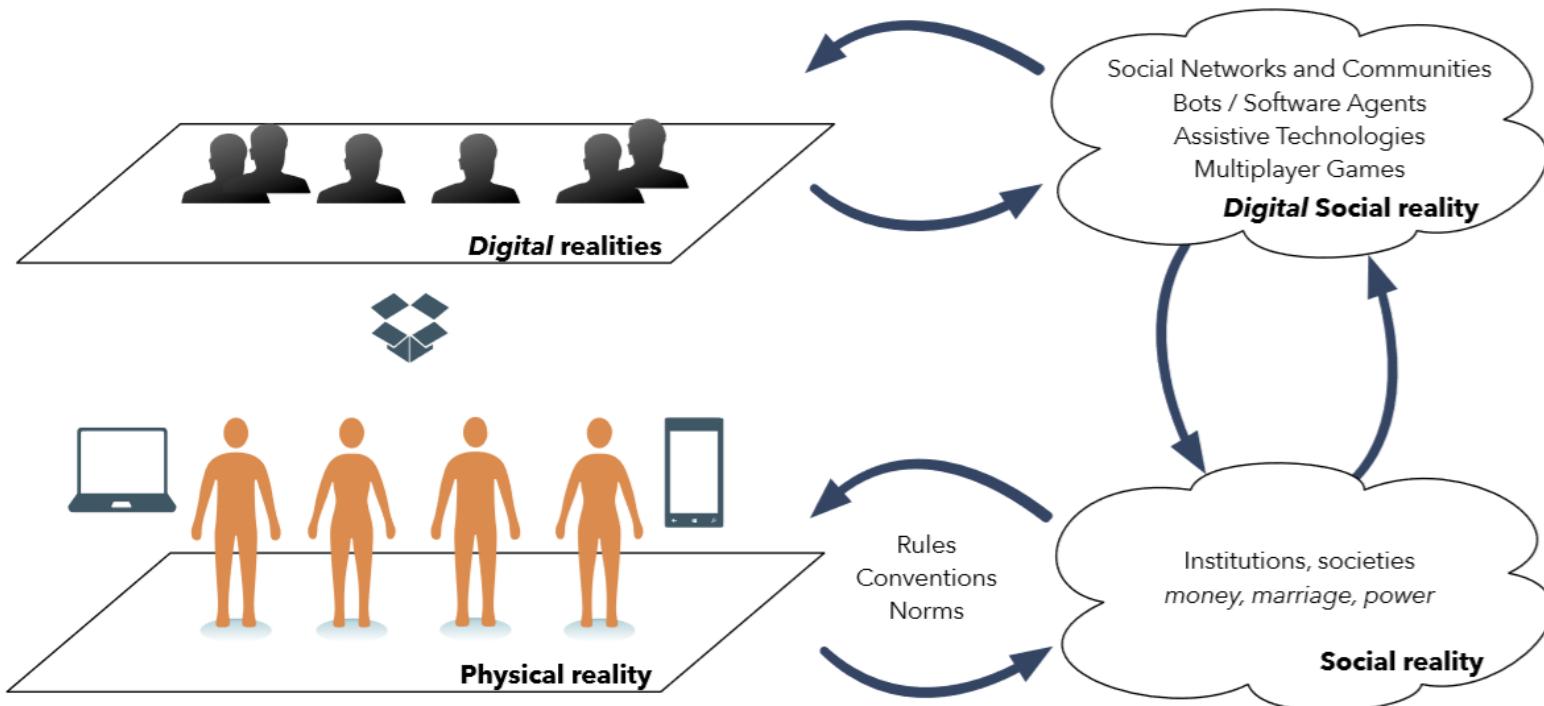
- *Realidad Social* (desde hace más de 20000 años)

Impacto en la sociedad: La IA en los Sistemas Sociotécnicos



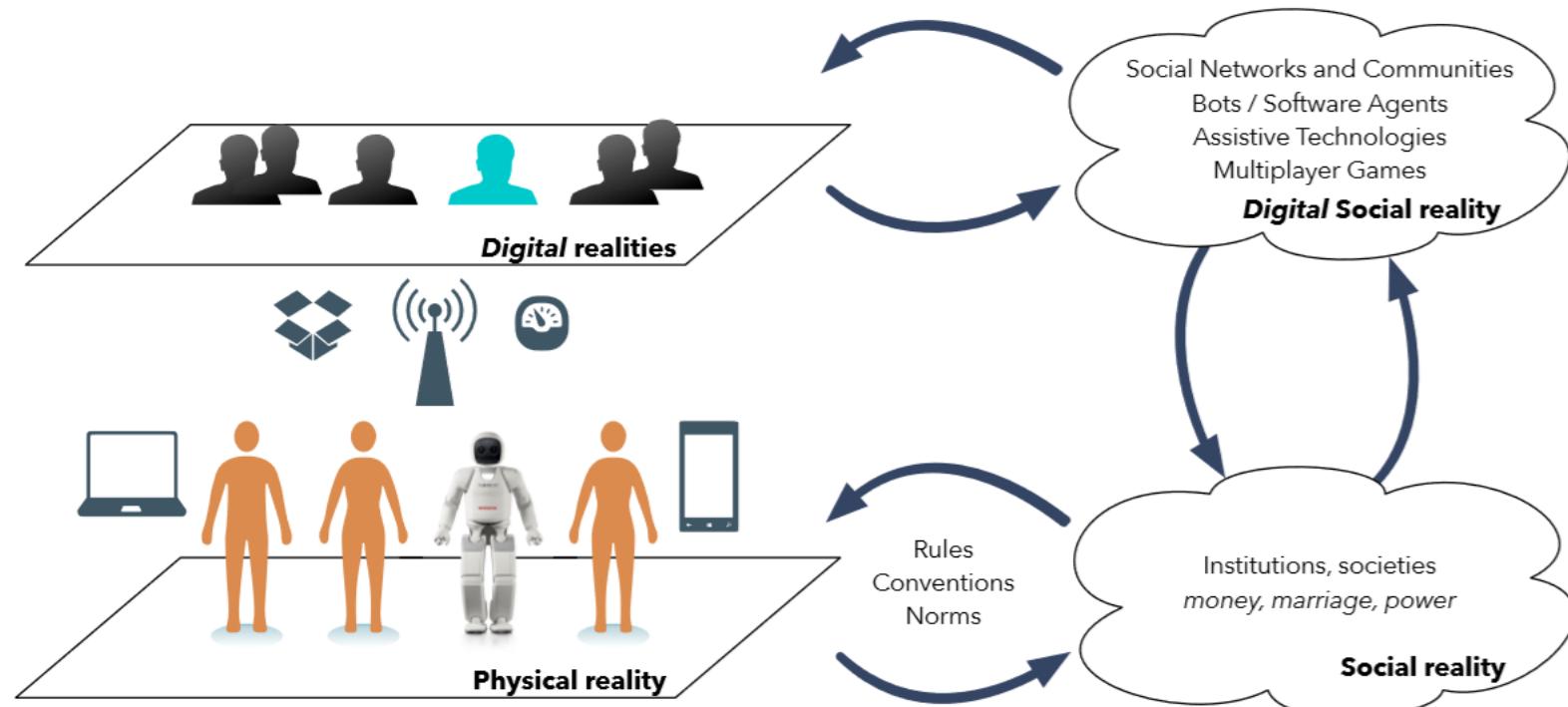
- Las TIC solo como facilitadoras de (casi lo mismo) de la *Realidad Social*

Impacto en la sociedad: La IA en los Sistemas Sociotécnicos



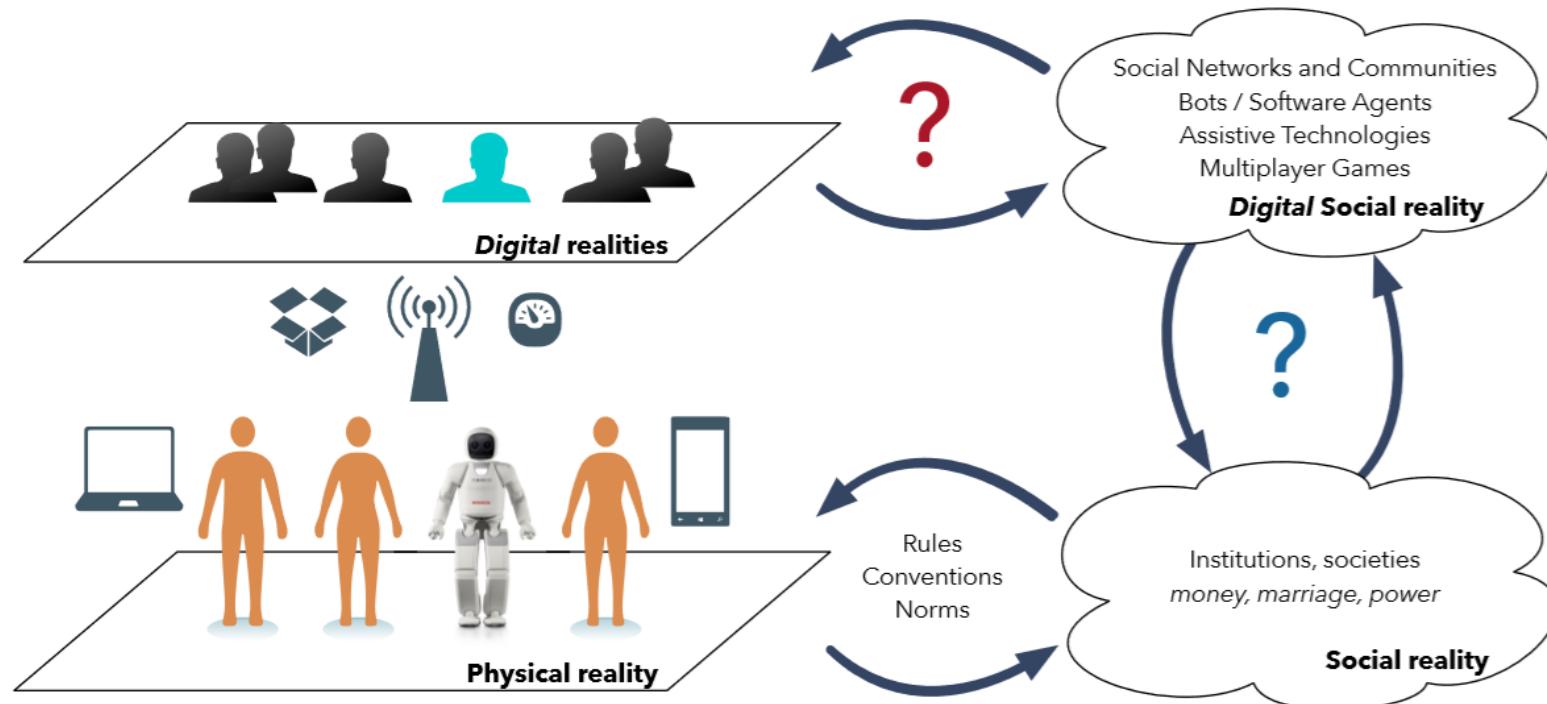
- Las TIC crean *nuevas realidades*
- Las TIC reconfiguran las relaciones humanas → los *sistemas sociotécnicos*

Impacto en la sociedad: La IA en los Sistemas Sociotécnicos



- Robots, Internet de las Cosas, Sensores → las TIC entran en *nuestra realidad física*

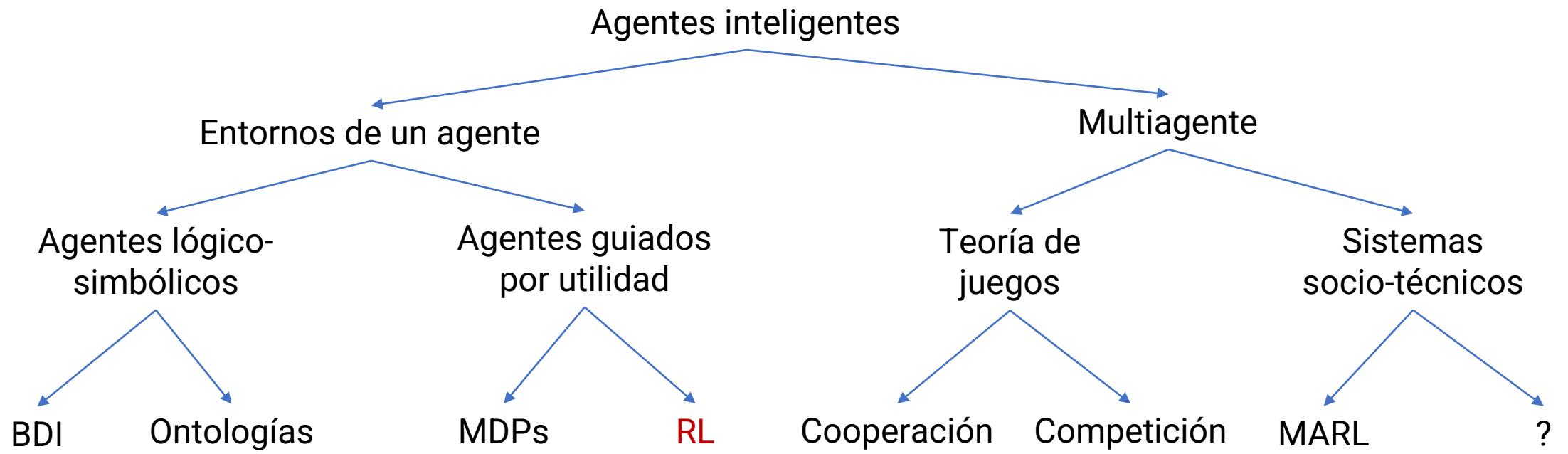
Impacto en la sociedad: La IA en los Sistemas Sociotécnicos



- Pregunta 1: Impacto de todas estas TIC en los seres humanos
- Pregunta 2: ¿Cómo se pueden aplicar los **aspectos sociales** a las entidades TIC?

Situándolo dentro del curso...

¿Qué hemos visto hasta ahora?



¿Podemos modelar de forma simbólica el comportamiento social esperado de un sistema de IA?

Coordinación

- Wooldridge y Jennings definen un Agente como un programa informático capaz de tomar sus propias decisiones sin control externo (**autonomía**), en función de sus percepciones del entorno y de los objetivos que pretende satisfacer. Un agente puede tomar acciones en respuesta a los cambios en el entorno (**reactividad**) y también puede tomar iniciativas (**proactividad**).
- Otro atributo de los agentes es su capacidad para comunicarse con otros agentes (**habilidad social**), no solo para compartir información sino, lo que es más importante, para coordinar acciones con el fin de alcanzar objetivos para los que los agentes no tienen planes que puedan cumplir por sí mismos, resolviendo problemas aún más complejos.

Coordinación

Definiciones

- **Coordinación** podría definirse como el proceso de gestión de las dependencias entre actividades. A través de este proceso, un agente razona sobre sus acciones locales y las acciones previstas que otros agentes pueden realizar, con el objetivo de hacer que la comunidad se comporte de manera coherente.
- Una **actividad** es un conjunto de operaciones potenciales que un **actor** (asumiendo un **rol**) puede realizar, con un objetivo determinado o un conjunto de objetivos.
- Un **actor** puede ser un **agente** o un grupo de agentes
- Un conjunto de **actividades** y una ordenación entre ellas es un **procedimiento**.

Coordinación

Modelos sociales de Coordinación

- Una fuente de inspiración para resolver problemas de coordinación son las **sociedades** humanas
- La **Sociología** es la rama de las ciencias que estudia las interrelaciones entre los individuos y la sociedad
- La **Teoría de la Organización** es un área específica dentro de la Sociología y la Economía que estudia la forma en que se pueden estructurar las relaciones en las organizaciones humanas (un tipo específico de sociedad)
- Son varias las abstracciones sociales que se han introducido en los Sistemas Multiagente
 - *Estructuras sociales y roles sociales*
 - *Organizaciones Electrónicas (organigrama)*
 - *Sistemas normativos (normas sociales)*.

Modelos Organizacionales

Modelos simbólicos de IA Social

Modelos sociales

Sociología y Sociedades

- **Sociología** es una disciplina que resulta de una evolución de la filosofía moral y ética con el fin de describir las interacciones que surgen entre los miembros de un grupo, y las estructuras sociales que se establecen.
- El objetivo de cualquier **sociedad** es permitir que sus miembros coexistan en un entorno compartido y persigan sus respectivos objetivos en presencia y/o en cooperación con otros.
 - Objetivos y requisitos globales
 - Previsibilidad
 - Reglas explícitas y posibilidades de interacción
- Esto también se puede aplicar a sociedades digitales compuestas por entidades computacionales
 - Sociedades de agentes

Modelos sociales

Rol (social)

- Uno de los principales conceptos que encontramos en las estructuras sociales complejas es el *rol*.
 - Un rol es una descripción de las tareas y objetivos a realizar por una entidad.
 - La idea es que no es importante quién desempeña el papel en la medida en que haya suficientes entidades que lo representen.
- Los roles han sido ampliamente estudiados en el campo de la *Teoría de la Organización*, con el fin de estudiar
 - las relaciones entre los roles sociales que un individuo puede desempeñar,
 - las obligaciones y autorizaciones que se asocian a cada una de esas funciones, y
 - la interacción de roles en la distribución de los mecanismos laborales.

Modelos sociales

Diseño organizacional

- El **Diseño organizacional** implica la creación de roles, procesos y relaciones formales de reporte en una organización.
- Se puede distinguir entre dos fases en un proceso de diseño organizacional:
 - *Agrupación estratégica*, que establece la estructura general de la organización (sus principales subunidades y sus relaciones), y
 - *Diseño operacional*, que define los roles y procesos más detallados.
- El libro más citado es Thompson (1967); otros trabajos clave incluyen Galbraith (1973) y Lawrence & Lorsch (1967).

Modelos sociales

Estructuras sociales

- En sistemas abiertos, se debe definir algún tipo de estructura para facilitar la coordinación en un escenario de control distribuido.
- Una buena opción tomada de las interacciones entre humanos y animales es la definición de las *estructuras sociales*.
- Las *estructuras sociales* definen un *nivel social* donde el sistema multiagente se considera como una *sociedad* de entidades con el fin de mejorar la coordinación de las actividades de los agentes (como la gestión del paso de mensajes y la asignación de tareas y recursos) mediante la definición de *patrones estructurados de comportamiento*.

Modelos sociales

Clasificación sociológica

- Según Findler *et al.* las estructuras sociales se clasifican en:
 - Una **alianza** es un grupo temporal formado voluntariamente por agentes cuyos objetivos son bastante similares. Los agentes renuncian, mientras están en la alianza, a algunos de sus propios objetivos y cooperan plenamente con los otros miembros de la alianza. Los agentes permanecen en la alianza mientras sea de su interés, a partir de entonces pueden unirse a otra alianza o permanecer por su cuenta.
 - Un **equipo** es un grupo formado por un agente especial (llamado líder del equipo) que recluta miembros calificados para resolver un problema determinado.
 - Una **coalición** es similar a una alianza, ya que es un grupo temporal en el que los miembros no abandonan sus objetivos individuales, sino que se dedican sólo a aquellas actividades conjuntas cuyos objetivos no están en conflicto con sus propios objetivos.
 - Una **convención** es una descripción formal de objetivos o acciones prohibidas o preferidas en un grupo de agentes.
 - Un **mercado** es una estructura que define dos roles prominentes (**comprador** y **vendedor**) y define los mecanismos para realizar transacciones comerciales.

Modelos sociales

Clasificación organizacional

- V. Dignum propone un enfoque más genérico, donde las estructuras sociales se dividen en tres grupos:
 - **Mercado (Markets)**, donde los agentes están interesados en sí mismos, impulsados completamente por sus propios objetivos. La interacción en los mercados se produce a través de la comunicación y la negociación.
 - **Red (Network)**, donde las coaliciones de agentes interesados en sí mismos acuerdan colaborar para lograr un objetivo común. La coordinación se logra por interés mutuo, posiblemente utilizando terceros de confianza.
 - **Jerarquía (Hierarchy)**, donde los agentes son (casi) totalmente cooperativos, y la coordinación se logra a través de cadenas de mando y control.
- los tres grupos propuestos por V. Dignum tienen como objetivo clasificar tanto a las organizaciones humanas como a las de agentes software.

Modelos sociales

Clasificación organizacional

	Market	Network	Hierarchy
<i>Type of society</i>	Open	Trust	Closed
<i>Coordination</i>	Price mechanism	Collaboration	Supervision
<i>Relation form</i>	Competition	Mutual Interest	Authority
<i>Tone or climate</i>	Precision/ suspicion	Open-ended/ mutual benefits	Formal/ bureaucratic
<i>Conflict Resolution</i>	Haggling (Resort to courts)	Reciprocity (Reputation)	Supervision

Table 1.1: V. Dignum *et al.* classification of Social Structures.

- Esta clasificación es útil en la etapa de diseño, ya que trata de motivar la elección de una de dichas estructuras en función de su idoneidad para un entorno específico.

Modelos sociales

Abstracciones sociales: Roles

- Los **roles** identifican las actividades y servicios necesarios para alcanzar los objetivos sociales y permiten abstraerse de los individuos específicos que eventualmente los realizarán.
 - Desde la perspectiva del *diseño de la sociedad*, los roles proporcionan los bloques de construcción para los sistemas de agentes que pueden desempeñar el papel,
 - Desde la perspectiva del *diseño del agente*, los roles especifican las expectativas de la sociedad con respecto a la actividad del agente en la sociedad.
- En esencia, los modelos a seguir se ocupan de la colaboración y la coordinación y especifican las relaciones de colaboración entre entidades sin fijar a priori todo el proceso de interacción.

Modelos sociales

Abstracciones sociales: Dependencia entre roles

- La **dependencia entre roles** significa que un rol depende de otro rol para la realización de sus objetivos.
 - Las sociedades establecen *dependencias* y las *relaciones de poder* entre los roles, lo que indica las relaciones entre los roles.
 - Estas relaciones describen cómo los actores pueden interactuar y contribuir a la realización de los objetivos de los demás. Es decir, un objetivo de un rol se puede *delegar* o *solicitar* a otros roles.

Modelos sociales

Modelos de roles y modelos de interacción

- los **modelos de roles** reflejan la competencia social de los agentes
 - modelados por derechos y obligaciones
 - Influyen en el comportamiento de los agentes
 - dando lugar a interacciones (protocolos de comunicación) típicos para la construcción de la sociedad
- Los modelos a seguir permiten garantizar algunas características globales del sistema y, al mismo tiempo, preservar la flexibilidad individual
 - Los derechos y obligaciones explícitos permiten comprometerse con roles específicos
 - Los roles garantizan un comportamiento global
 - Las descripciones de roles están representadas por modelos formales
- los **modelos de interacción** reflejan los flujos de trabajo y los procesos de negocio
 - Procedimientos explícitos y acceso
 - Las descripciones de las escenas se especifican formalmente, lo que permite la verificación
 - Animación de sociedades

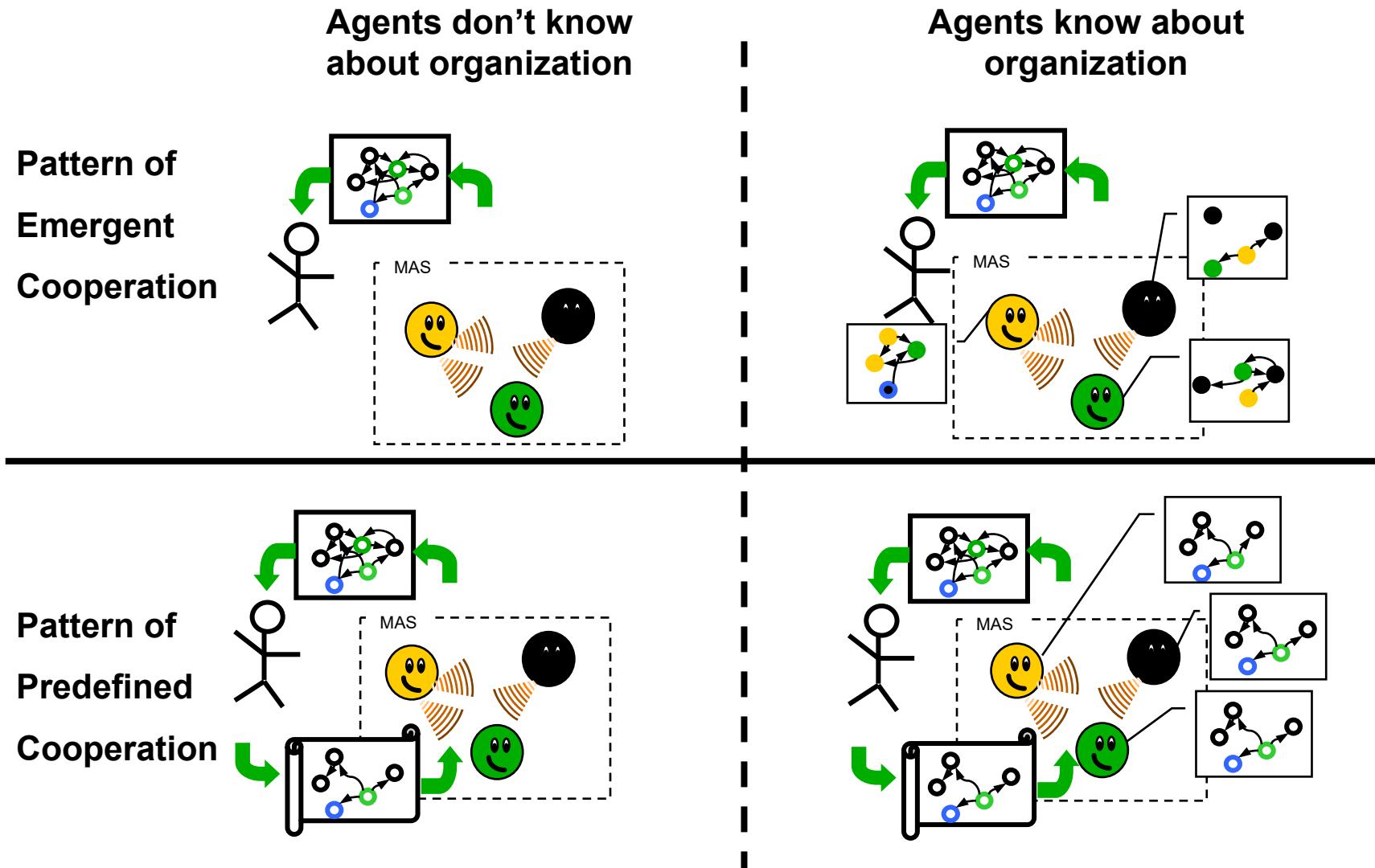
Organizaciones Electrónicas

Definición

- Una **Organización** es un patrón supra-agente de cooperación **emergente** o cooperación **predefinida** de los agentes en el sistema, que puede ser definida por el diseñador o por los propios agentes, para un propósito.
 - Patrón de cooperación emergente/potencial
 - *Entidad organizacional, relaciones sociales, compromisos, coaliciones*
 - Patrón de cooperación predefinido
 - *Estructura organizativa, normas, ...*

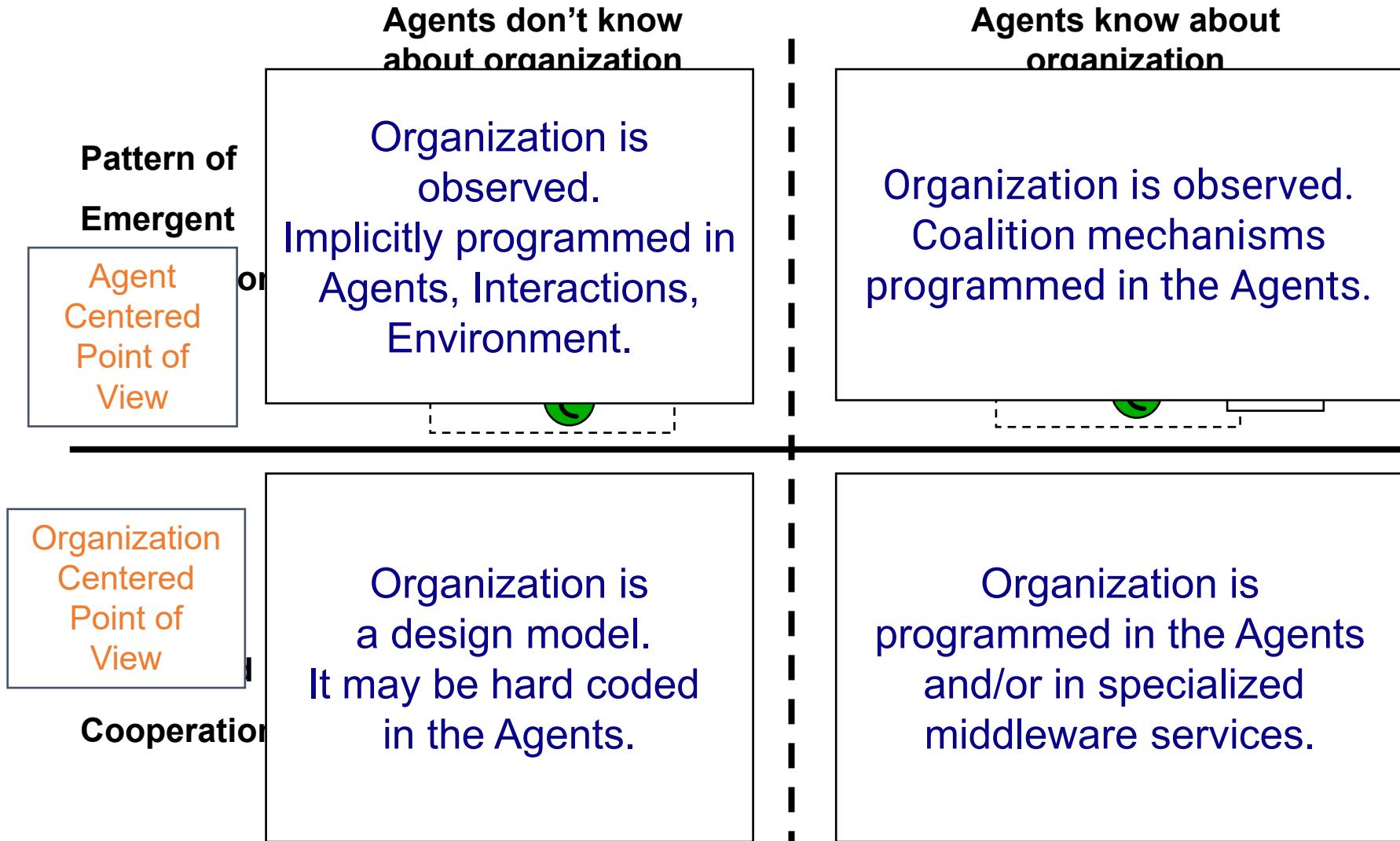
Organizaciones Electrónicas

cuatro perspectivas



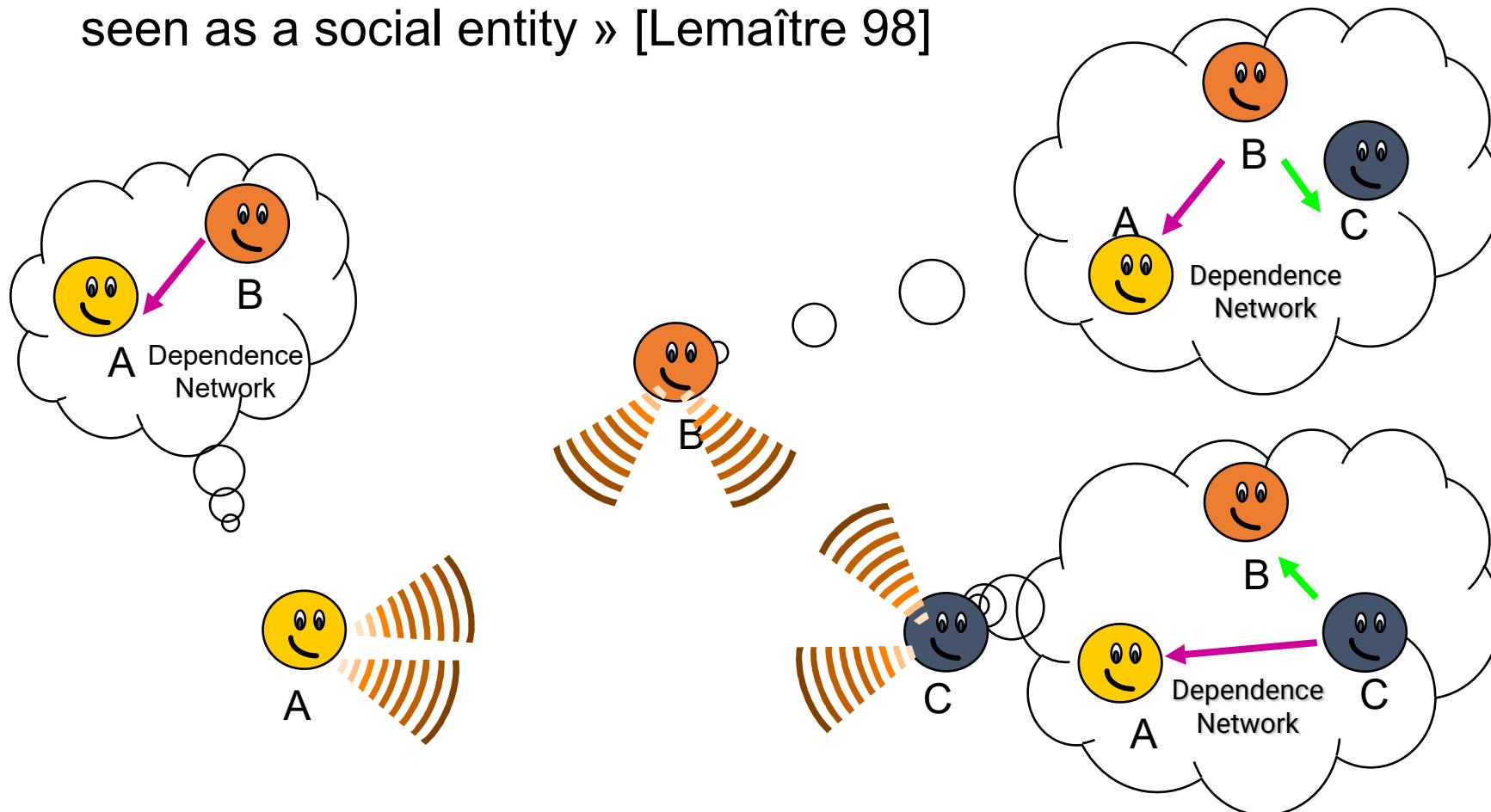
Organizaciones Electrónicas

¿Donde se programa la Organización?



Organizaciones Electrónicas perspectivas centradas en el agente

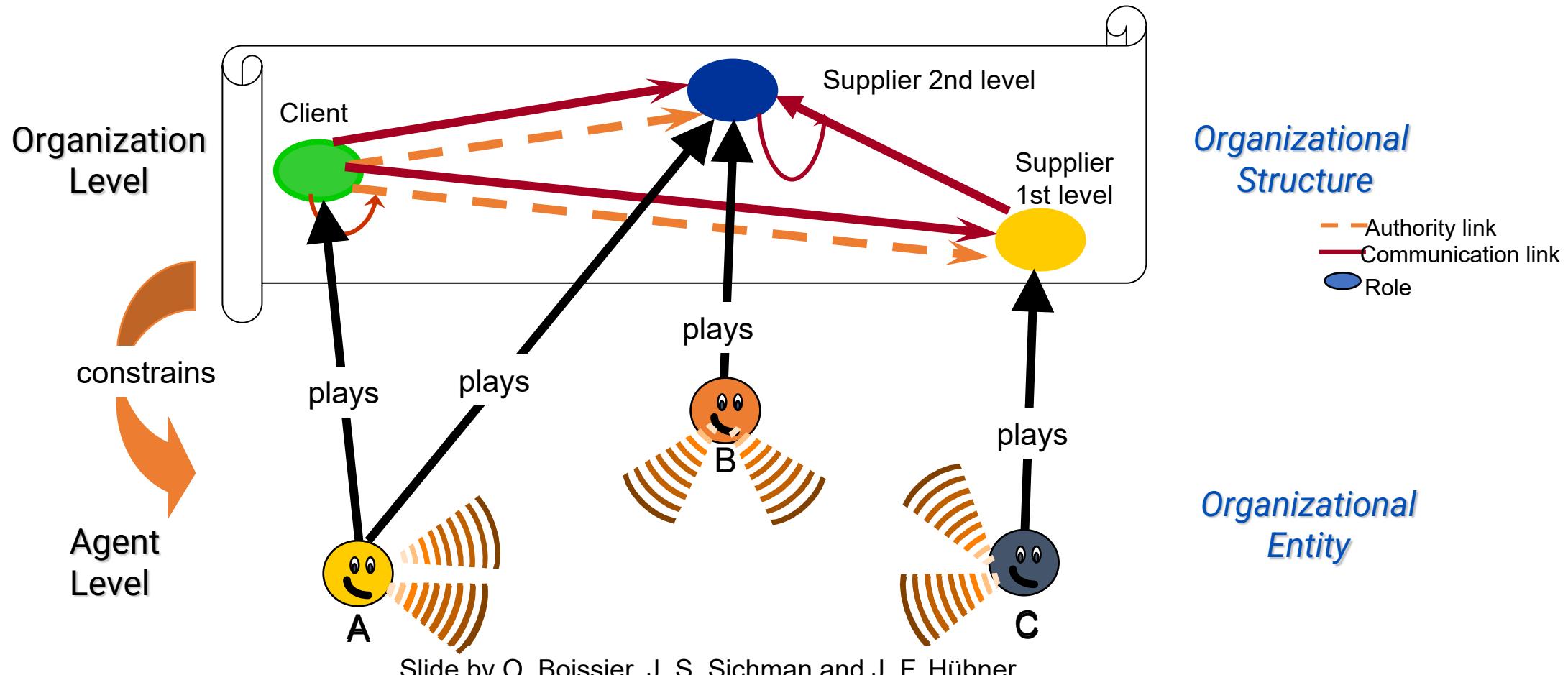
« The social concepts are all focused on the **agents'** behavior seen as a social entity » [Lemaître 98]



Organizaciones Electrónicas

perspectivas centradas en la organización

« The leading concept is the **group** or the **organization** instead of the **agent** » [Lemaître 98]



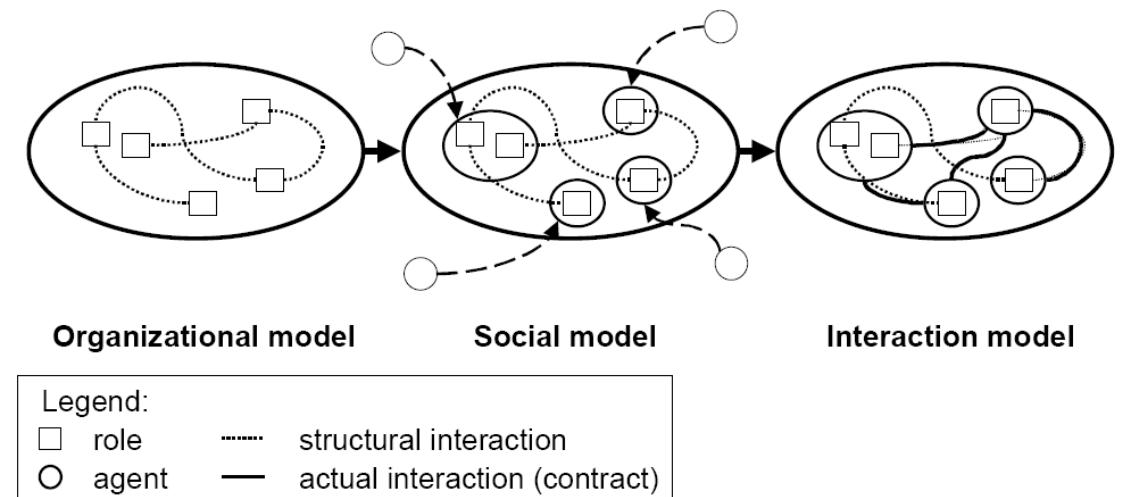
Organizaciones Electrónicas: OperA

Ejemplo: el OperA framework

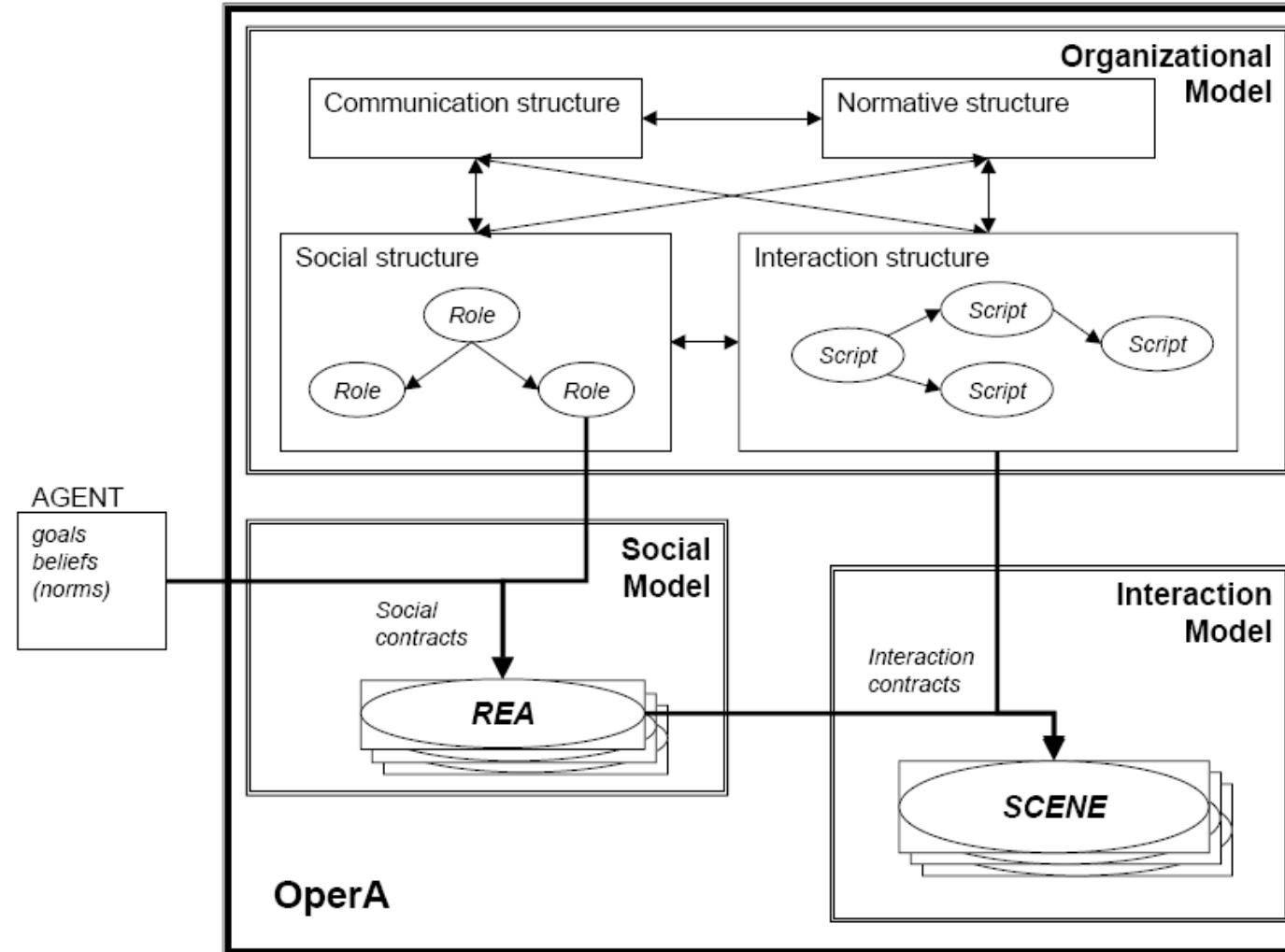
- OperA es un marco para la especificación de sistemas multiagente.
- Distingue entre:
 - los mecanismos a través de los cuales se describe y coordina la estructura y el comportamiento global del modelo, y
 - los objetivos y el comportamiento de los proveedores de servicios (agentes) que pueblan el modelo
- Se basa en una semántica formal que hace posible la verificación.
- OperA representa la interacción entre agentes de una manera que:
 - es independiente del diseño interno de los agentes
 - distingue las características de la organización de los objetivos propios de los agentes
 - Crea vínculos dinámicos entre el diseño organizacional y las poblaciones de agentes
 - permite la adaptación de los patrones de interacción a las características de poblaciones específicas.

Organizaciones Electrónicas: OperA

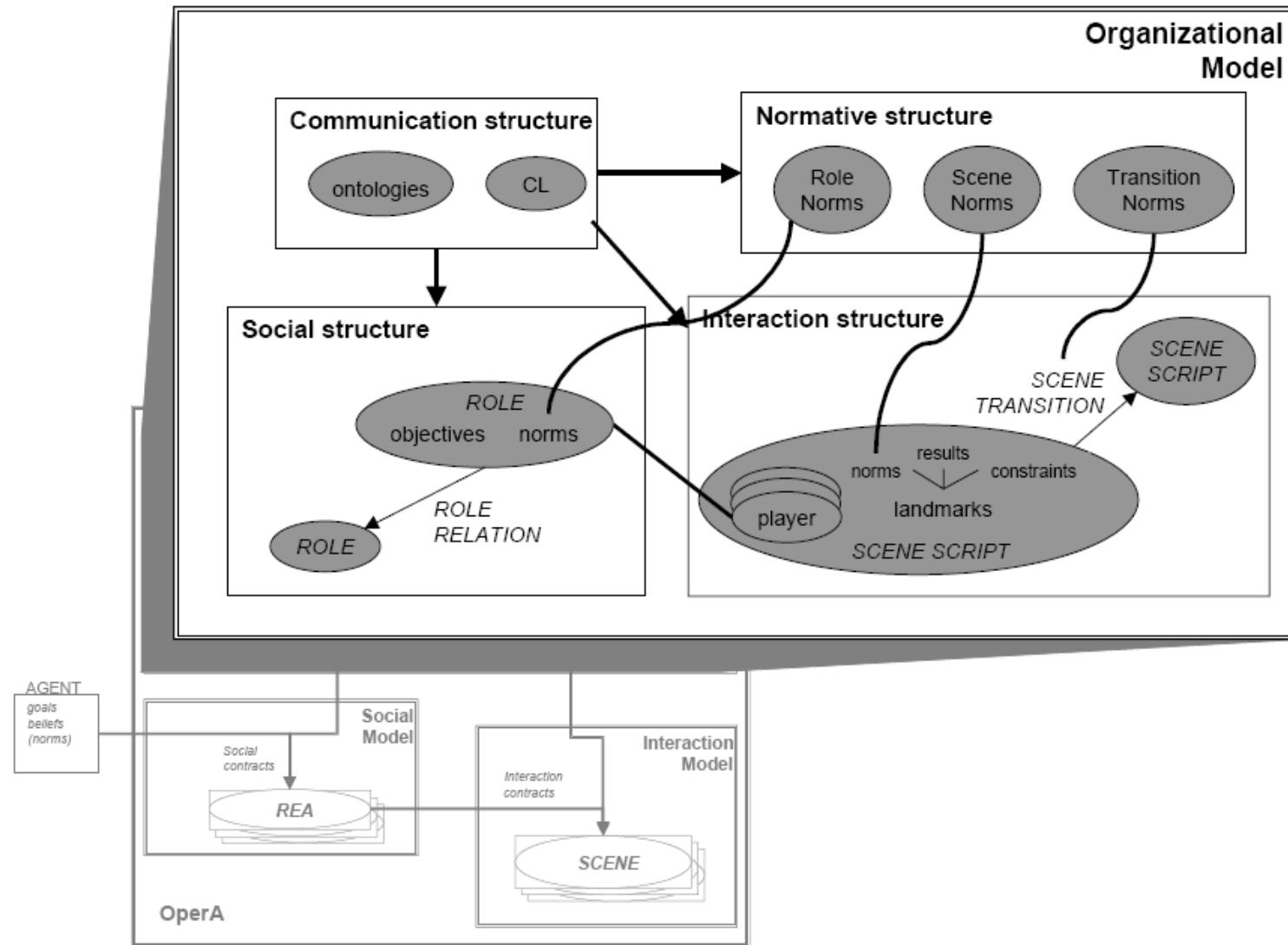
- 3 modelos:
 - **Organizational Model**
 - representa los objetivos y requisitos de la organización
 - modelo de roles, estructuras de interacción, guiones de escena, normas
 - **Social Model**
 - representa acuerdos relativos a la participación de agentes individuales
 - rea = role enacting agent
 - **Interaction Model**
 - acuerdos relativos a la interacción entre los propios agentes



Organizaciones Electrónicas: OperA



Organizaciones Electrónicas: OperA



Organizaciones Electrónicas: OperA

pasos del diseño del Organizational Model (OM)

- Identificar
 - partes interesadas (stakeholders): internas y externas
 - Tipos de coordinación
 - Requisitos y objetivos (globales y por parte interesada)
 - Dependencias de roles
 - Normas
- Describir ontologías de dominio
 - Conceptos (y posibles relaciones)
- Especificar la estructura social
- Identificar el orden parcial de los objetivos y dependencias de los roles
- Describir guiones de escenas
- Identificar los requisitos y consecuencias de la transición de la escena
- Especificar la estructura de interacción

Organizaciones Electrónicas: OperA

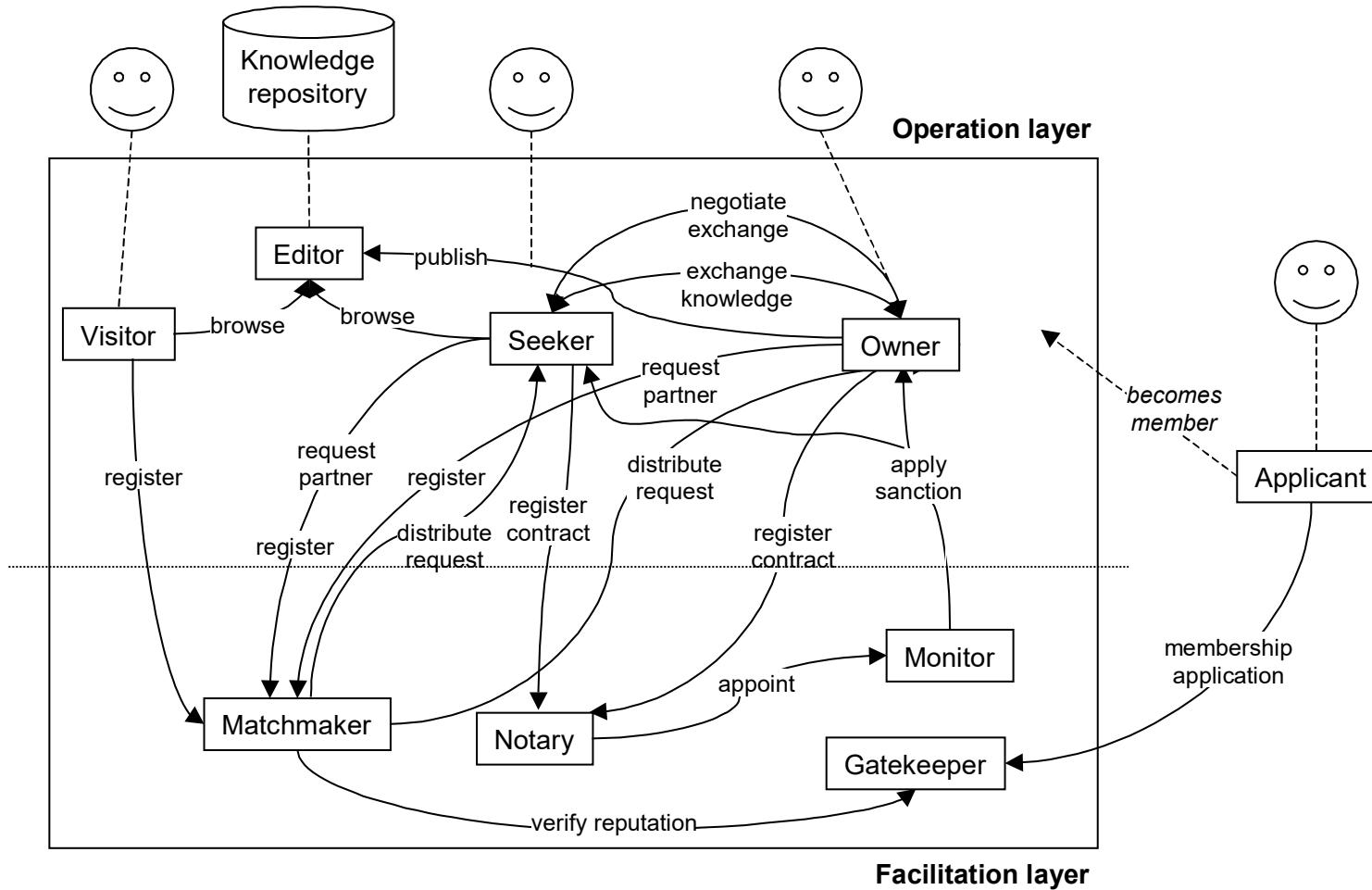
OM: Social Structure

- Definición de Roles
- Los subobjetivos se identifican mediante el análisis de medios y fines
- Las dependencias de rol identifican → necesidades de coordinación
→ interacciones necesarias (a definir en el IM)

ROLE DEFINITION	
Role id	Identified in Environment Level
Objectives	Formalization of objectives identified in the role table
Sub-objectives	Result of means-end analysis for each role objective
Rights	From means-end analysis and norm analysis
Norms	From the Norm analysis in Environment Level
Type	Roles associated with the coordination model are institutional, and operational roles are in principle external.

Organizaciones Electrónicas: OperA

OM: definir dependencias entre roles



Organizaciones Electrónicas: OperA

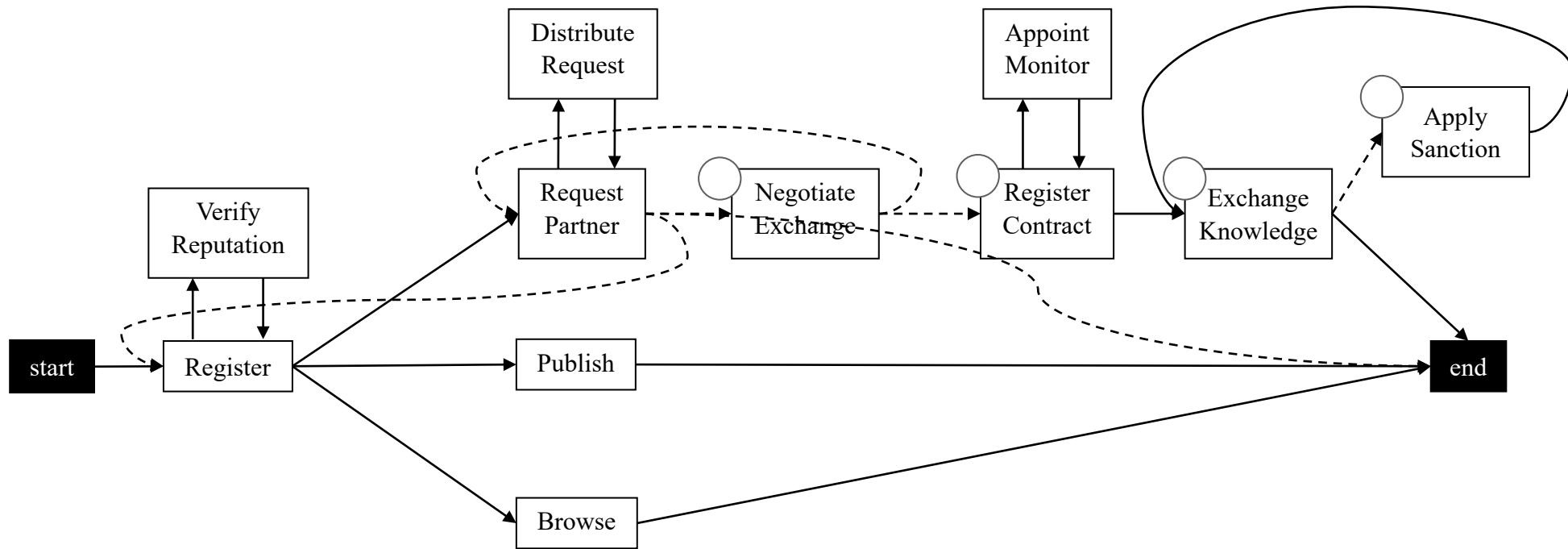
OM: Normative Structure

- Identificar las normas de la organización
 - Análisis de responsabilidad
 - Análisis de recursos
 - Análisis de disparadores
 - Especificación de la norma
 - Especificación de la sanción
- Norm tables

Description	Norm Analysis	
Type (O, F, P) <i>Textual description</i>	Responsibilities	Initiator: <i>role</i> Action: <i>role</i>
	Triggers	Pre condition: <i>state</i> Post condition: <i>state</i>
	Specification	whenever state then role is O,F,P to achieve state otherwise sanction

Organizaciones Electrónicas: OperA

OM: Interaction Structure

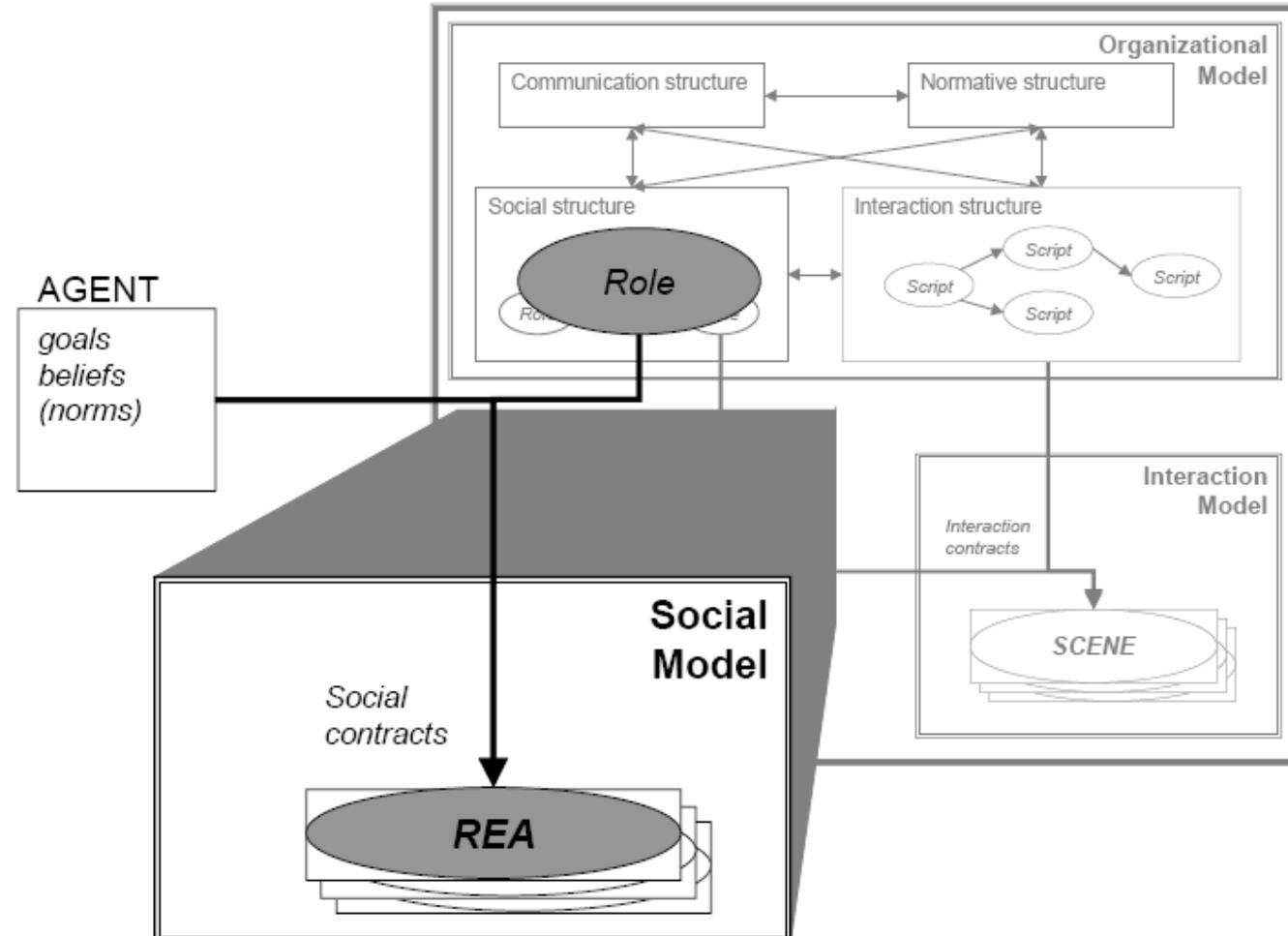


Organizaciones Electrónicas: OperA

OM: Interaction Structure - scene script

Interaction Scene: Partner Request	
Description	Seeker requests possible partners that can answer knowledge need
Roles	S: Knowledge-seeker(1), M: Matchmaker (1)
Results	DONE receive-partners(S, M, question, ListPartners)
Patterns	<pre> graph LR A[Request partners] --> B[Distr request] B --> C[Distr dd] C --> D[answer request] C --> E[Receive partners] D --> F[Answer dd] F --> G[req dd] </pre> <p style="text-align: right;">landmarks!</p> <p>BEFORE answer-deadline)}</p> <p>}</p>
Norms	<p>OBLIGED request-knowledge(M, knowledge-owners, answer-deadline) BEFORE deadline</p> <p>IF request-knowledge(matchmaker, P, question, deadline) THEN OBLIGED answer-request(P, M, YN, question) BEFORE deadline</p>

Organizaciones Electrónicas: OperA



Organizaciones Electrónicas: OperA

pasos del diseño del Social Model (SM)

- Especificación de **role negotiation scenes**
- Especificación del Negotiator Agent
- Basado en las **Descripciones de roles** especificadas en el OM
 - Requisitos mínimos
 - características negociables y su alcance

Role Negotiation Scene: Role R	
Roles	Negotiator (N), applicant (A)
Results	$\rho = \text{contract}(A, R, \text{SocialContract})$
Plans	$\Pi\rho = \{ \text{agreed(char1)} \text{ AND } \dots \text{ AND } \text{agreed(charN)} \text{ BEFORE } \text{contract-agreed}(N, A, \text{social-contract}(A, R, CC)) \}$
Norms	PERMITTED(N, negotiate-social-contract(A, R)). OBLIGED (N, propose-range(char1, min, MAX))...

Organizaciones Electrónicas: OperA

Contratos

- Un **contrato** es una declaración de intenciones que regula el comportamiento entre organizaciones e individuos
 - Normas específicas
 - Períodos de tiempo (*inicio, fin, deadlines*)
 - Términos y condiciones
 - Sanciones
- Enfoque de los contratos en OperA
 - **Roles a desempeñar** (social contracts)
 - Desarrollo de las **escenas** (interaction contracts)

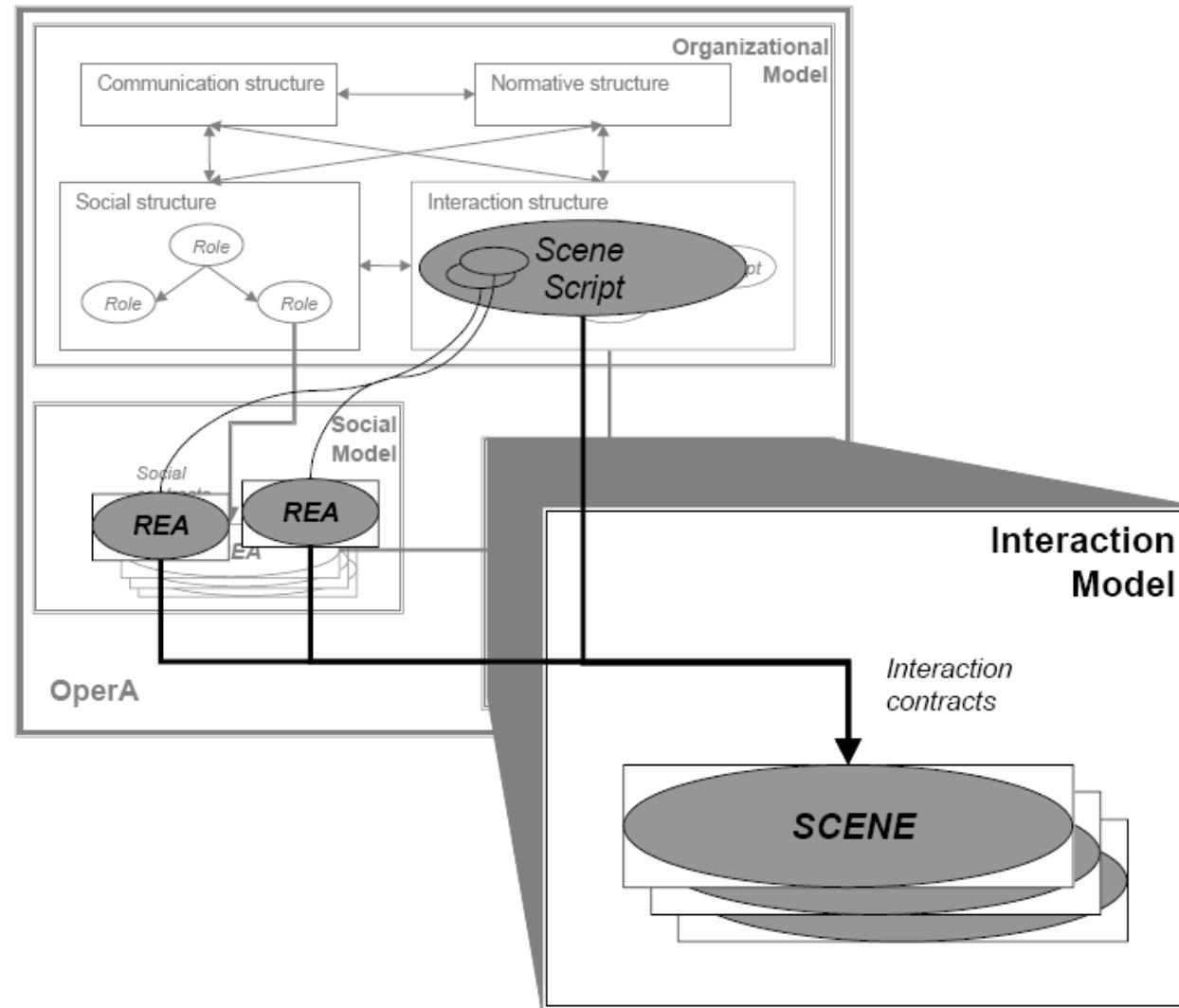
Organizaciones Electrónicas: OperA

SM: Social Contracts

- El Social Model genera y gestiona (en tiempo de ejecución) instancias de:
 - Role enacting agent
 - Social contract

Social Contract	
Agent	Anne
Role	Knowledge seeker
Clauses	
	<ol style="list-style-type: none">1. PERMITTED(Anne, access-kb([KB1, KB3, KB7])2. OBLIGED(Anne, publish-received-knowledge(item, KB3) allows(KO, publish))3. $\forall p: \text{contract}(p, \text{Anne}) \rightarrow \text{PERMITTED}(p, \text{publish}(p, \text{Anne's-item}, kb))$

Organizaciones Electrónicas: OperA

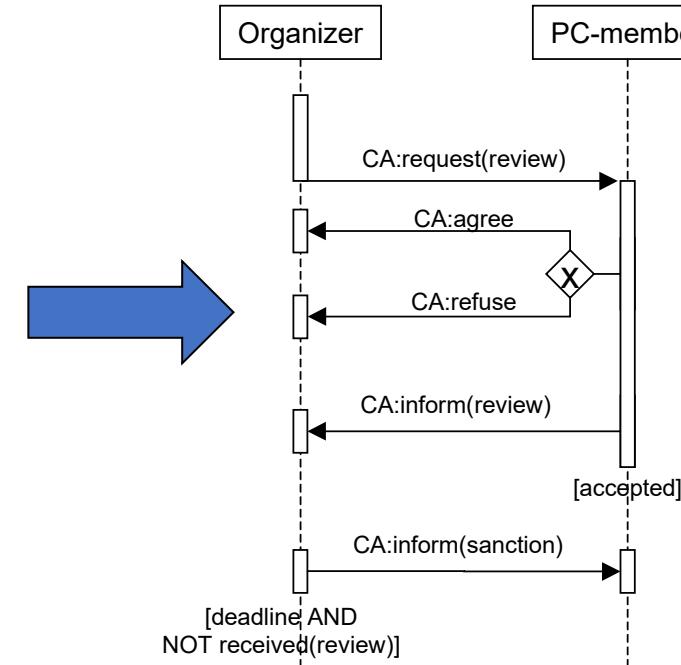


Organizaciones Electrónicas: OperA

IM: Interaction Contracts

- Basado en la definición previa de [script interaction scenes](#)
- Depende de las características de los agentes que se postulan a los roles de la sociedad
- Resultado de los Interaction Contracts:
 - [Protocolos específicos](#) para las escenas, que son soportados por los agentes

Interaction Scene: PC-member role enactment	
Roles	Society keeper (SK), applicant (A), society register (R)
Results	$\rho = \text{contract}(A, \text{PCmember}, \text{SocialContract})$
Plans	$\Pi_p = \{\text{agreed}(\text{max-papers}(M)) \text{ AND } \text{agreed}(\text{review-deadline}(D)) \text{ BEFORE }$ $\text{contract-agreed}(\text{SK}, A, \text{social-contract}(A, \text{PCmember}, CC)) \text{ BEFORE }$ $\text{contract-registered}(R, \text{social-contract}(A, \text{PCmember}, CC))\}$
Norms	PERMITTED(SK, negotiate-social-contract(A, PCmember)).
	OBLIGED (SK, role-description-announced(role(PCmember))).



Modelos Normativos

Modelos simbólicos de IA Social

Modelos Normativos

Instituciones como Estructuras sociales

- Las **Estructuras sociales** definen un nivel social que potencia la coordinación por medio de *patrones de interacción*
- Las **Instituciones** son un tipo de estructura social en la que un corpus de **normas** moldea el comportamiento de los miembros de un grupo
- La definición de una Institución (humana) suele incluir
 - Reglamentos sobre las interacciones
 - Convenciones: ‘hechos institucionales’ vs ‘hechos brutos’
 - ej. 1 : ‘asesinato’ (vs homicidio)
 - ej. 2 : ‘incesto’ (vs acto sexual)
 - ej. 3 : ‘propiedad’ (vs posesión física)
 - ej. 4 : ‘matrimonio’ (vs vivir juntos)
 - Procedimientos y protocolos para la creación y determinación de hechos institucionales

Modelos Normativos

Normas en e-Instituciones

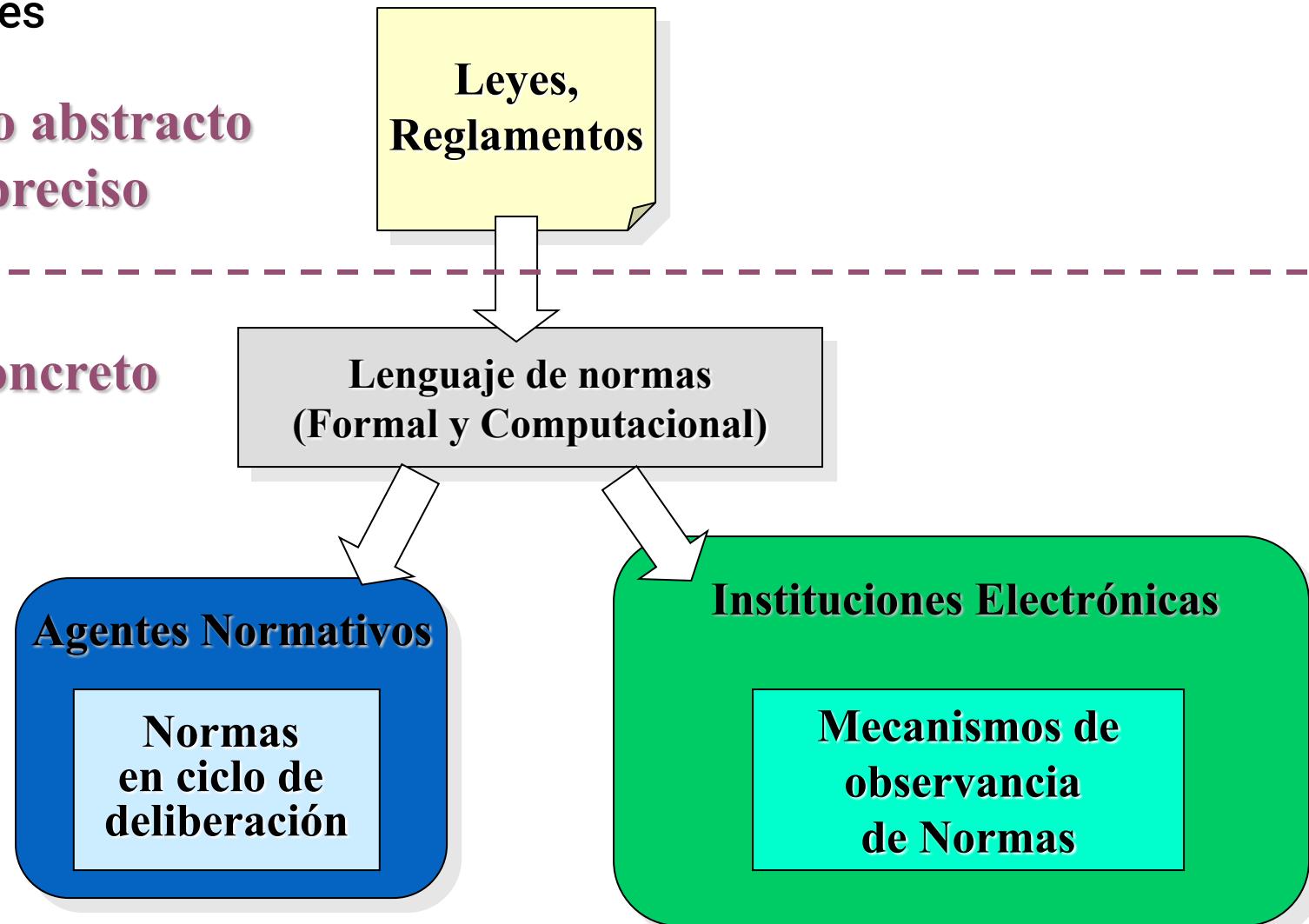
- Una **e-Institución** es el modelo computacional de una institución a través de
 - La especificación de las **normas** de la institución en (algun) formalismo computacional adecuado.
 - La especificación de los procedimientos y protocolos de la institución
 - En el contexto del MAS:
 - Reduce la **incertidumbre** del comportamiento de otros agentes
 - Reduce los **malentendidos** en la interacción
 - Permite a los agentes **prever** el resultado de una interacción
 - **Simplifica** la toma de decisiones (reducir las posibles acciones)
- **comportamiento** del Agente **guiado por las Normas**

Modelos Normativos

Elementos principales

**demasiado abstracto
e impreciso**

más concreto

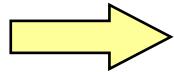


Modelos Normativos

representación de Normas

- Representación formal de las normas es necesaria
- ¿Qué lógica?

- Las normas permiten, obligan o prohíben
- Las normas pueden ser condicionales
- Las normas pueden tener aspectos temporales
- Las normas son relativas a los roles



variante de la Lógica Deontica

- La representación debe ser fácilmente parseable y utilizable por los agentes

OBLIGED, PERMITTED, FORBIDDEN
IF C
BEFORE D, AFTER D

Modelos Normativos

representación de Normas

- Normas incondicionales sobre predicados

- las normas sobre el valor de P están activas en todo momento:

OBLIGED(a, P) PERMITTED(a, P) FORBIDDEN(a, P)

- Normas incondicionales sobre acciones

- las normas sobre la ejecución de A están activas en todo momento:

PERMITTED($a \text{ DO } A$) FORBIDDEN($a \text{ DO } A$)

- Normas condicionales

- La activación de las normas está condicionada a C

- C puede ser un predicado sobre el sistema o el estado de una acción:

OBLIGED($(a, P) \text{ IF } C$)
PERMITTED($(a, P) \text{ IF } C$)
FORBIDDEN($(a, P) \text{ IF } C$)

OBLIGED($(a \text{ DO } A) \text{ IF } C$)
PERMITTED($(a \text{ DO } A) \text{ IF } C$)
FORBIDDEN($(a \text{ DO } A) \text{ IF } C$)

Modelos Normativos

representación de Normas

- Normas condicionales con Deadlines
 - La activación/violación de las normas está definida por una fecha límite

OBLIGED((*a, P*) BEFORE *D*)

PERMITTED((*a DO A*) AFTER *D*)

FORBIDDEN((*a, P*) BEFORE *D*)

- Deadlines absolutos y relativos:

23:59:00 09/05/2004

time(done(assign(organ, recipient))) + 5min

- Ejemplos:

OBLIGED((*allocator DO assign(heart, recipient)*)
BEFORE (*time(done(extraction(heart, donor))) + 6hours*))

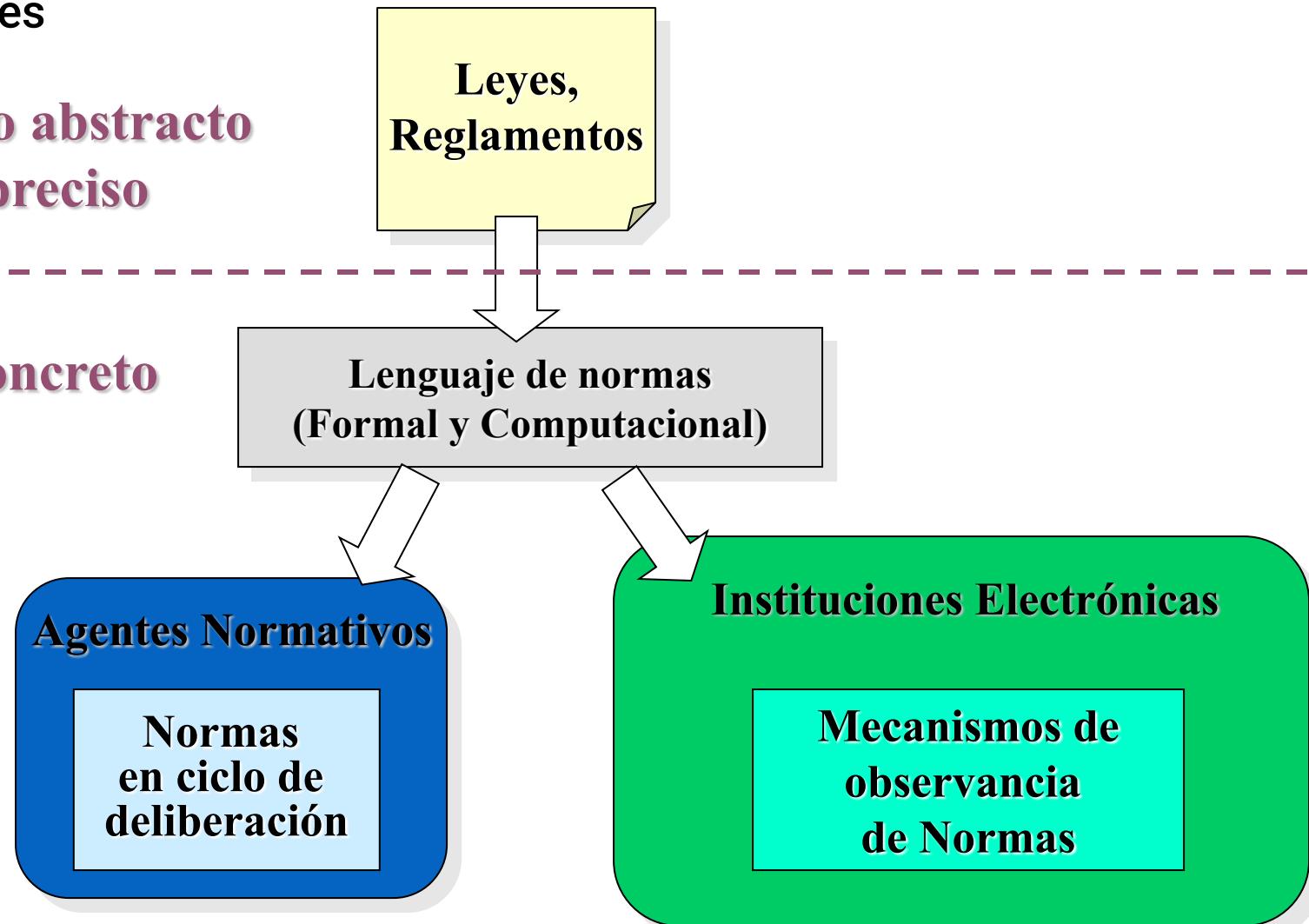
FORBIDDEN((*allocator DO assign(organ, recipient)*)
IF NOT(*hospital DONE ensure_quality(organ)*))

Modelos Normativos

Elementos principales

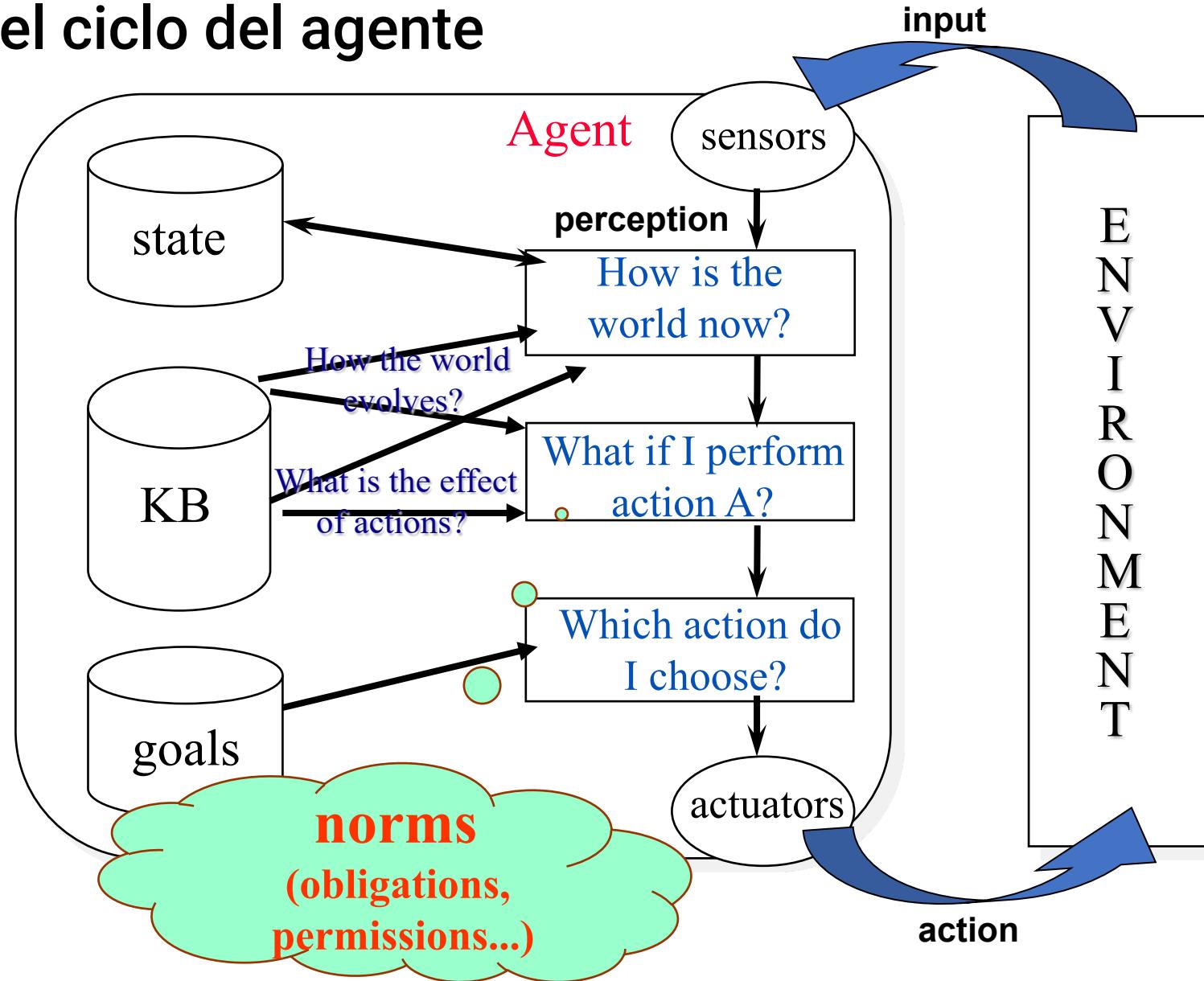
**demasiado abstracto
e impreciso**

más concreto



Agentes Normativos

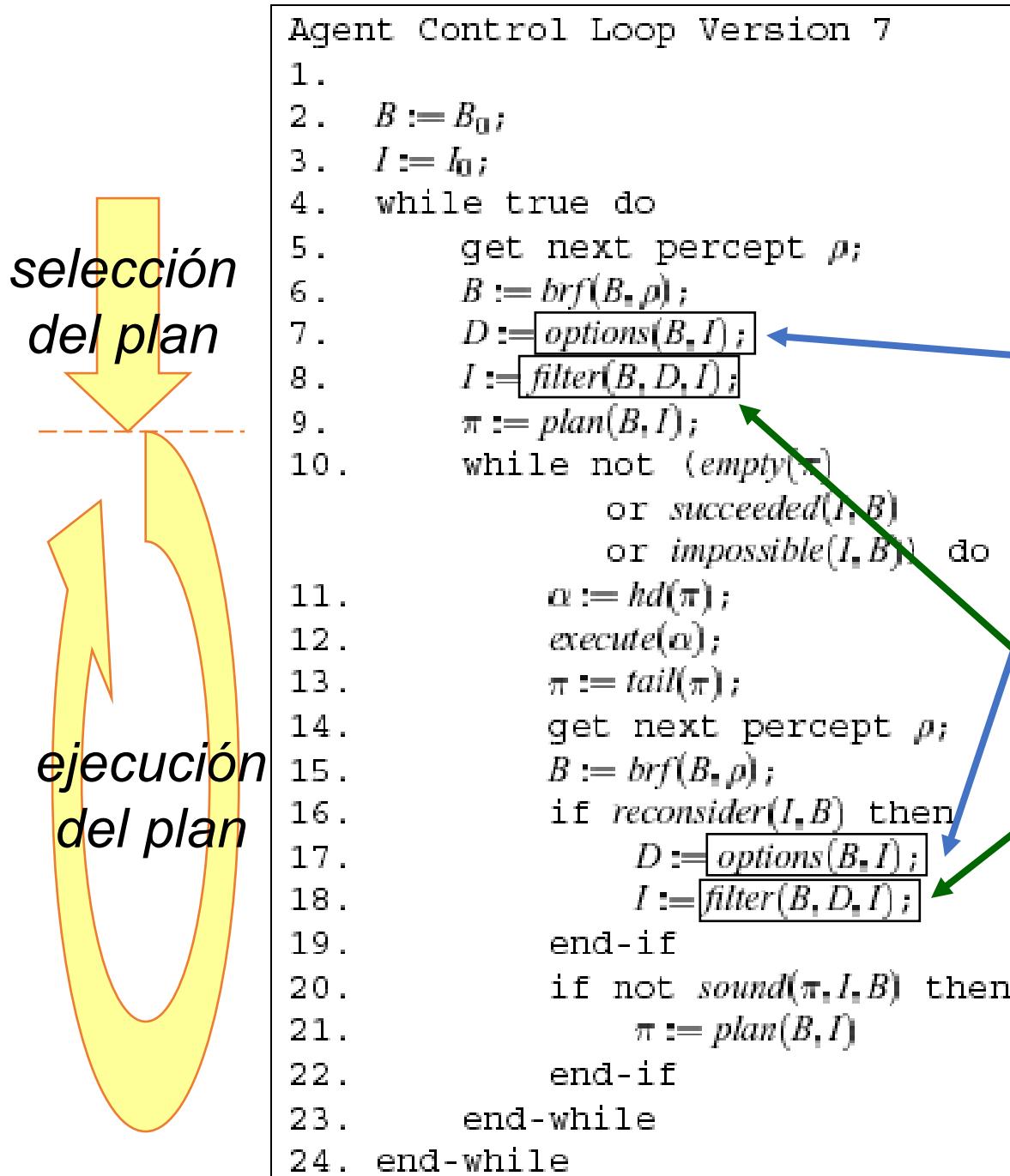
Normas en el ciclo del agente



Agentes Normativos

Normas en el ciclo del agente

- ¿Como las normas influencian el comportamiento del agente?
 - El agente no tiene conocimiento de las normas
 - Las normas están integradas en el código del agente
 - Las normas están integradas en los planes y protocolos que utiliza el agente
 - Las normas son elementos explícitos en el razonamiento del agente
 - El agente puede o no adoptar las normas
 - El agente puede o no seguir las normas
 - El agente sigue la norma siempre que sea posible
 - El agente viola la norma a veces
 - El agente viola la norma siempre que sea posible



Obligaciones normativas
añaden acciones al conjunto de opciones y puede definir algunas prioridades o precedencias

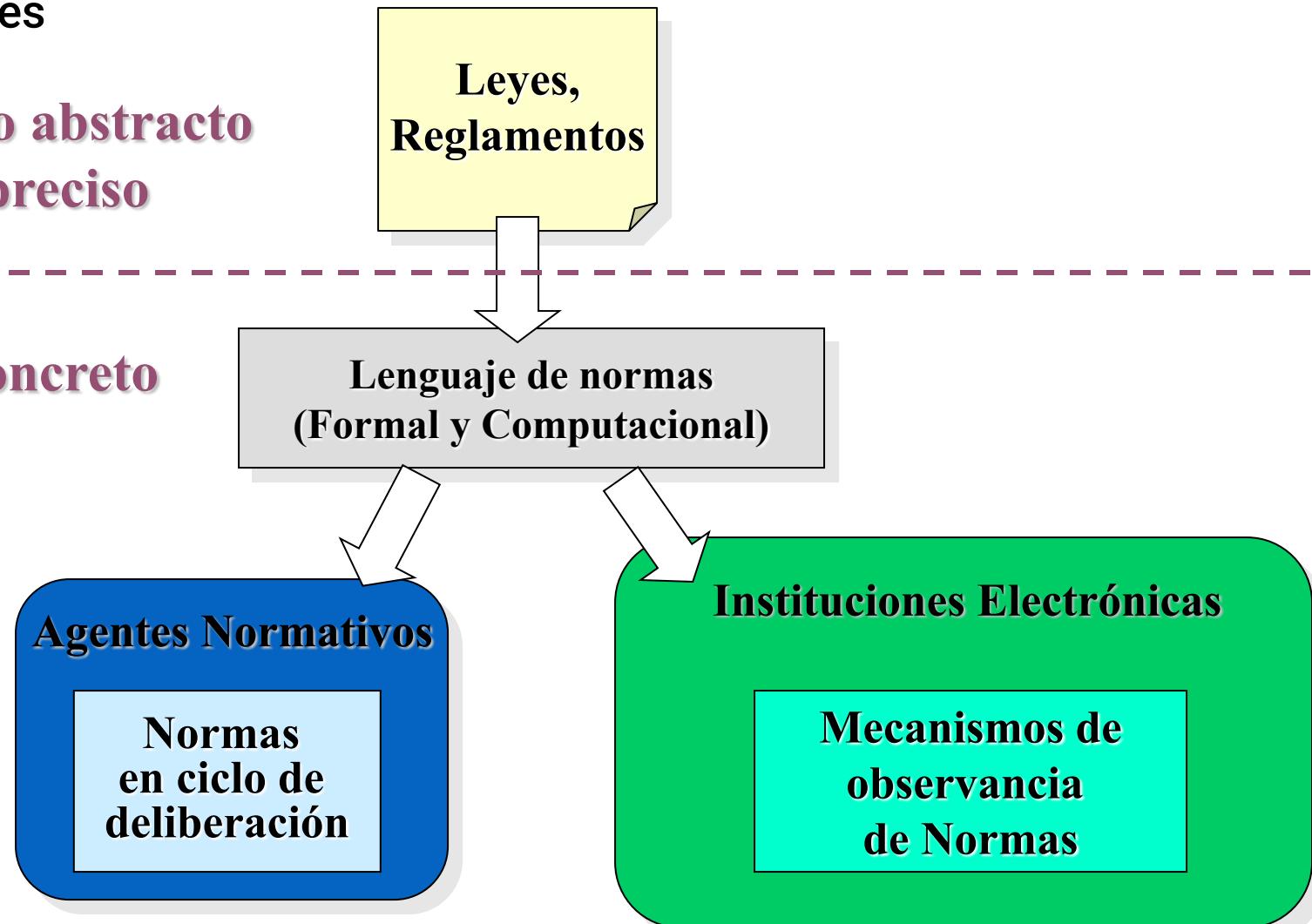
Prohibiciones normativas
eliminan acciones del conjunto de opciones

Modelos Normativos

Elementos principales

**demasiado abstracto
e impreciso**

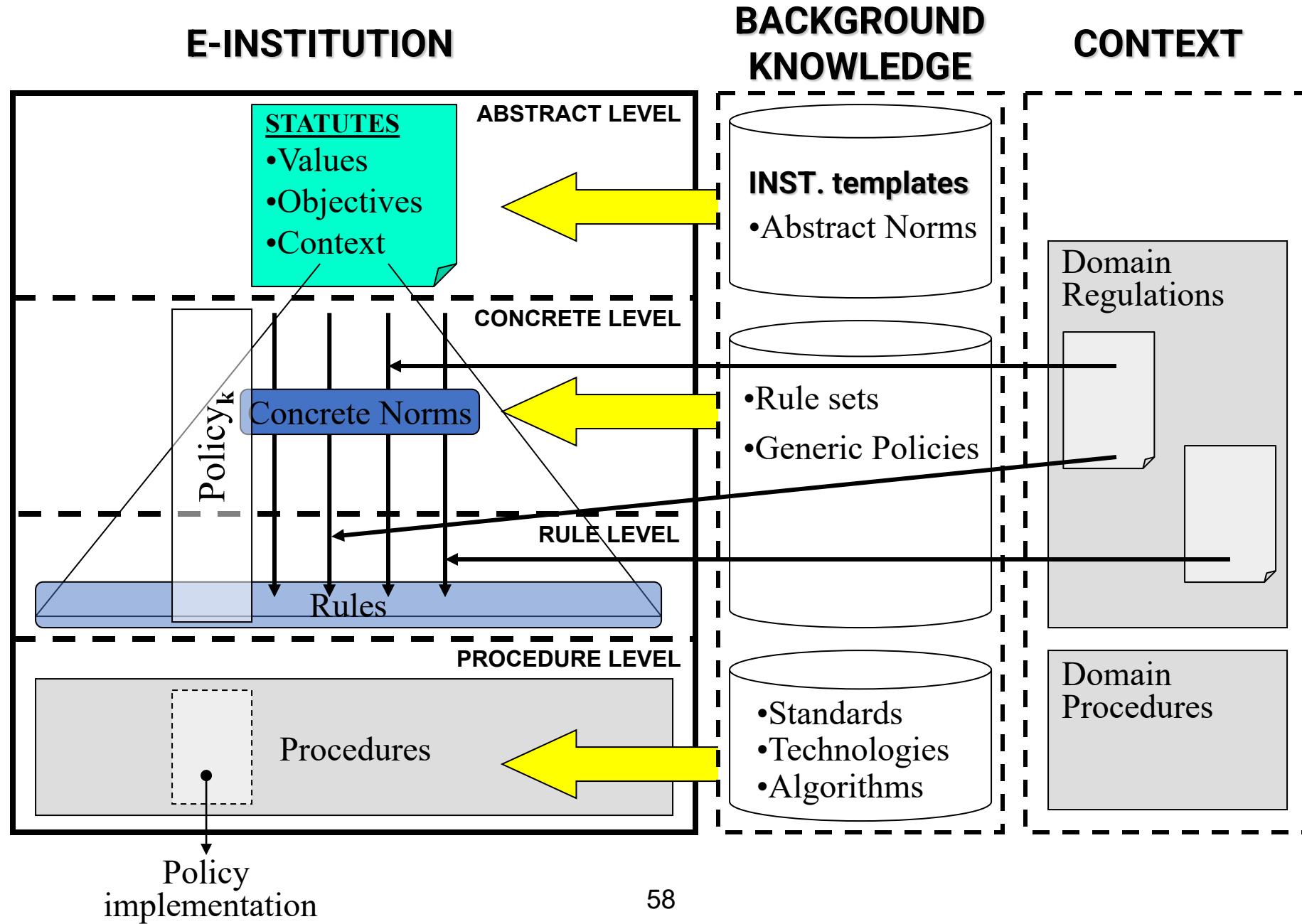
más concreto

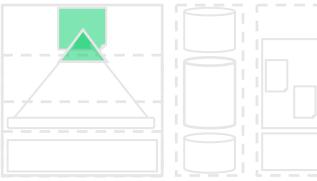


Entornos Normativos: e-Instituciones

HARMONIA

- HARMONIA es un marco multivel que propone una conexión formal entre los diferentes niveles de abstracción de un Sistema Normativo
 - Distingue entre niveles normativos y prácticos
 - Análisis del Contexto en los sistemas normativos
 - *contexto y normas, contextos anidados, influencia de los supercontextos*
 - Establece la conexión entre la especificación formal y la implementación del agente
 - top-down: guiando el diseño
 - bottom-up: seguimiento del origen de un protocolo/plan
 - Observancia de Normas como la detección de estados ilegales/no deseados





Entornos Normativos: e-Instituciones

HARMONIA - Abstract Level

- Los estatutos definen la
 - Valores
 - Objetivos
 - Contextode la organización.

E.g.: *Organización Nacional de Trasplantes:*

El objetivo principal de la ONT es aumentar el número de donaciones de órganos y el consiguiente aumento de los órganos disponibles para trasplantes.

La ONT funciona de acuerdo con la regulación del Sistema Nacional de Salud y se esfuerza por distribuir los órganos donados de la manera más adecuada y correcta de acuerdo con los conocimientos técnicos actuales y de acuerdo con los principios éticos de igualdad.



Entornos Normativos: e-Instituciones

HARMONIA - Abstract Level

- Los **Valores** son creencias que tenemos sobre lo que es importante y, por lo tanto, sobre cómo deberían ser las cosas.
 - *"Distribución adecuada"*
 - *"Distribución según principios éticos de igualdad"*
 - *"Equidad de la transacción"*
 - *"Respetar la privacidad de las personas"*
- Los **Valores** pueden ser considerados como el nivel más abstracto en el que se expresan las normas.
- Los valores de una organización pueden definirse mediante el conjunto de **Normas Abstractas** especificadas en la org. que contribuya a ese valor

```
D(equity) := {FONT (discriminate(x,y,age)),  
              OONT (find_best_recipient(organ)),  
              }  
D(appropriate(distribution)) := OONT(appropriate(distribution))
```



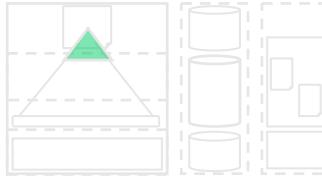
Entornos Normativos: e-Instituciones

HARMONIA - Abstract Level

- **Problema:** Estas normas son demasiado **abstractas**, ya que utilizan conceptos que no están completamente descritos en la ontología del agente, de la organización o del entorno de ejecución.
 - *"It is forbidden to discriminate based on age"*

$$F_{\text{ONT}}(\text{discriminate}(x,y,\text{age})),$$

- Las normas pueden ser **abstractas** de las siguientes maneras:
 - Se refieren a una **acción abstracta**
 - Usan **términos que son vagos**
 - Hacen abstracción de los aspectos **temporales**
 - Hacen abstracción de los **agentes y/o roles**
 - Se refieren a acciones o situaciones que no son (directamente) **controlables y/o verificables** por la e-institución



Entornos Normativos: e-Instituciones

HARMONIA - Abstract Level

- Ejemplo 1: Acciones abstractas

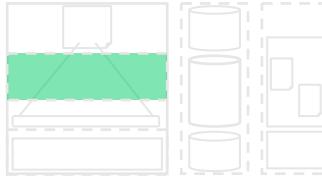
“a living donor should consent to the donation of an organ”

$$\begin{aligned} \text{sign(donor,contract)} \cup \text{carry(donor,will)} \cup \\ \text{tell(donor,family)} \end{aligned} \quad \left. \right\} \Rightarrow_{\text{ONT}} \text{Consent(donor)}$$

- Ejemplo 2: Términos vagos

“the ONT is obliged to ensure that the distribution of organs and tissues is appropriate”

$$\begin{aligned} \text{O}_{\text{ONT}}(\text{ensure_quality(organ)}) \cup \\ \text{O}_{\text{ONT}}(\text{ensure_compatibility(organ, recipient)}) \end{aligned} \quad \left. \right\} \Rightarrow_{\text{ONT}} \text{O}_{\text{ONT}}(\text{appropriate(distribution)})$$



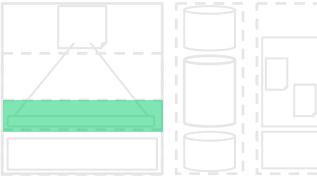
Entornos Normativos: e-Instituciones

HARMONIA - Concrete Level

- Las **normas concretas** son el resultado de traducir las normas abstractas en el contexto de la organización en normas que hacen uso de términos y acciones que se definen en la ontología de la organización.

$$O_{hosp}(\text{consent}(\text{donor}(p,x)) < \text{done}(\text{transplant}(hosp,x,p,q)))$$

- **Problema:** ¿CÓMO se implementa una norma concreta como esta en una organización electrónica?



Entornos Normativos: e-Instituciones

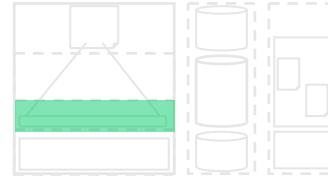
HARMONIA - Rule Level

- Traducción de la Dimensión Normativa a una Descriptiva
 - Idea: reducción de Deontic Logic a Dynamic Logic

$$O_{hosp}(\text{consent}(\text{donor}(p,x)) < \text{do}(\text{transplant}(\text{hosp},x,p,q)))$$

$$[\text{transplant}(\text{hosp},x,p,q)]\text{done}(\text{consent}(\text{donor}))$$
$$O_{buyer}(\text{pay}(\text{goods},\text{seller},\text{price}) < \text{do}(\text{exit}(\text{buyer})))$$

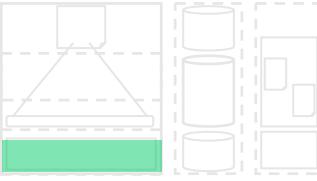
$$\text{not}(\text{done}(\text{pay}(\text{goods},\text{seller},\text{price}))) \rightarrow [\text{exit}(\text{buyer})]\text{V}(\text{fine}(\text{buyer}))$$



Entornos Normativos: e-Instituciones

HARMONIA - Rule Level

- Reglas, violaciones y sanciones
 - Las **reglas de violación** definen las infracciones
 - Una **violación** está compuesta por
 - pre-condicion
 - sanción
 - efectos laterales (side effects)
 - Las **Pre-condiciones** son utilizados por los **Agentes Policía** para detectar violaciones.
 - Las **Sanciones** son utilizadas por los **Agentes Normativos** para razonar sobre la utilidad de infringir la norma relacionada.
 - Los **Efectos laterales** son utilizados por los **Agentes Internos de la organización** para recuperar el sistema de la violación.



Entornos Normativos: e-Instituciones

HARMONIA - Rule Level

- Idea: la implementación final del sistema
- Formalmente, la traducción de **Dynamic Logic** a un **Lenguaje Procedural**:

$$\{P\} a \{Q\} \equiv P \rightarrow [a]Q$$

- Ejemplo:

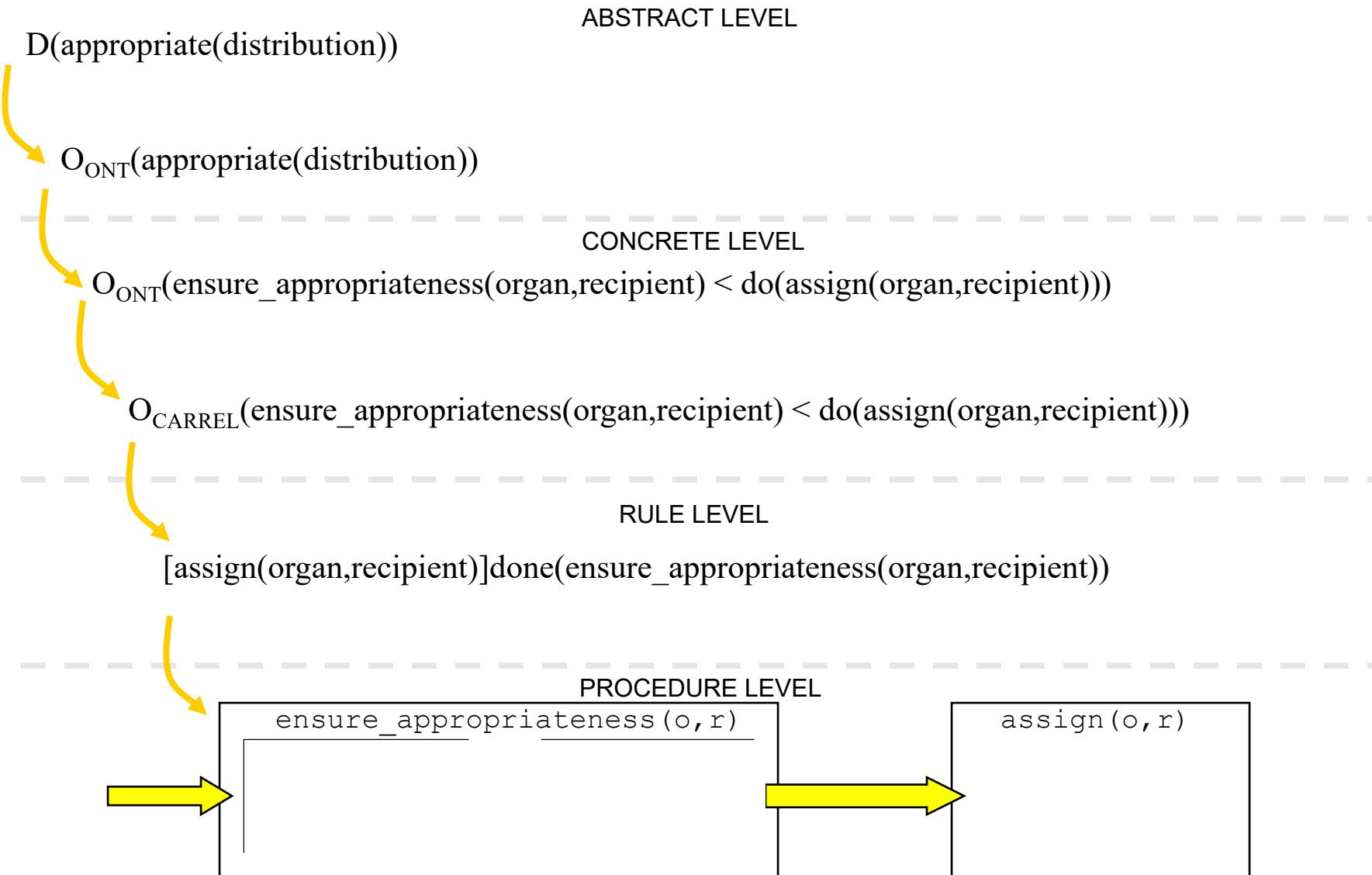
(not(done(assign(o, r))) \wedge done(ensure_appropriateness(o,r)))
 $\rightarrow [assign(o,r)]$ done(assign(o, r)



{not(done(assign(o, r))) & done(ensure_appropriateness(o,r))}
assign(o, r)
{done(assign(o, r))}



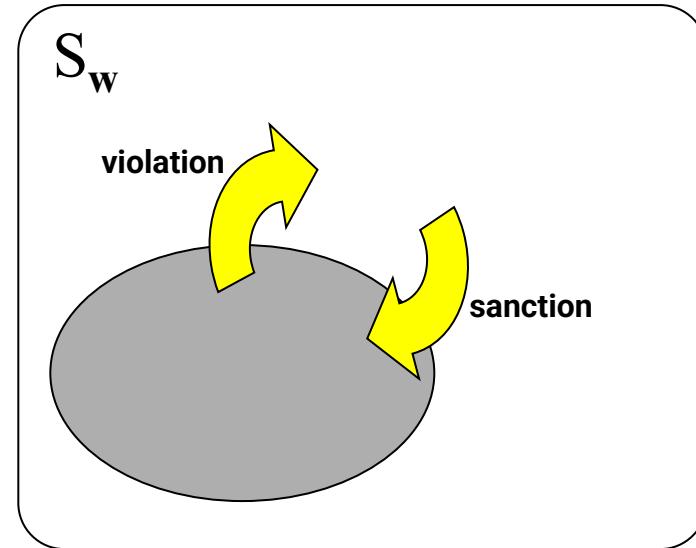
HARMONIA – Conexión entre niveles



HARMONIA – Normas como base de la seguridad y robustez

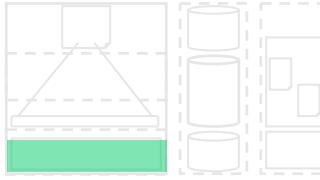
- En **HARMONIA** las normas definen los límites de un comportamiento aceptable

- Comportamiento deseado (legal) y no deseado (ilegal)
- Estados aceptables (seguros) e inaceptables (inseguros)



- Las **Violaciones** ocurren cuando los agentes rompen una o más normas, ingresando en un estado ilegal (inseguro).
- Las **Sanciones** son acciones para hacer que los agentes vuelvan a ser legales (seguros).
- Las sanciones incluyen las acciones para recuperar el sistema de una infracción

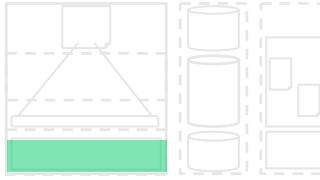
} Robustez
} Seguridad



Entornos Normativos: e-Instituciones

HARMONIA - Procedure Level

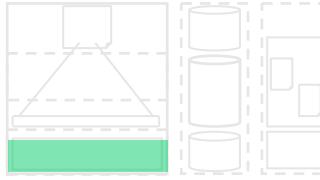
- Implementación de normas desde una perspectiva institutional perspective Implementación de un demostrador ≠ de teoremas para comprobar el cumplimiento del protocolo
- Implementación de un entorno seguro (**observancia de normas**)
- 2 opciones en función del control sobre los agentes
 - Definición de restricciones en el comportamiento no deseado
 - Definición de las infracciones y reacción ante ellas
- Las asunciones de **HARMONIA**:
 - A veces, las normas pueden ser violadas por los agentes
 - El estado interno de los agentes no es observable ni controlable
 - Las acciones no pueden imponerse sobre las intenciones de un agente
 - agentes como cajas negras
 - solo su comportamiento y acciones observables



Entornos Normativos: e-Instituciones

HARMONIA - Procedure Level

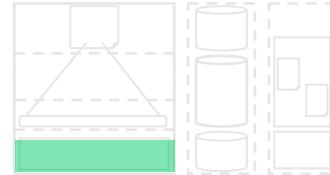
- La observancia de las normas no está centralizada, sino que se distribuye en un conjunto de agentes, los Agentes Policía: verifican si una acción determinada (*observable*) fue legal o ilegal dadas las condiciones de violación definidas para ese contexto.
- La Plataforma de Agentes debe ayudar a los Agentes de Policía, proporcionando ayudas rápidas y muy eficientes para la aplicación de las normas como servicios y mecanismos adicionales de la plataforma.
- *A) Detección de la ocurrencia de una acción*
 - Los Agentes Policía pueden sobrecargarse revisando TODAS las acciones
 - **black list mechanism** (de acciones a monitorizar) ej.: *assign*
 - **action alarm mechanism** (alarma a los Agentes Policía)
 - El Agente de Policía comprueba si se aplican las condiciones de una violación.



Entornos Normativos: e-Instituciones

HARMONIA - Procedure Level

- *B) Detección de activación/desactivación de normas*
 - activación = cuando la condición C es verdadera
 - desactivación = cuando P se mantiene, A está hecho o C es falso
 - tiempo de reacción: Tiempo permitido entre la activación de la norma y la reacción (antes de considerarlo violación)
 - Dependiendo de la complejidad a comprobar C, la plataforma debe implementar las **estructuras de datos de acceso rápido** adecuadas y/o **mecanismos de procesamiento** para reducir la carga de cómputo de los Agentes Policía
- *C) Control de deadlines*
 - **clock trigger mechanism**: genera eventos a una cierta hora o pasado un intervalo de tiempo para detectar que se ha superado un deadline



Entornos Normativos: e-Instituciones

HARMONIA - Procedure Level: ejemplo de definición de violación

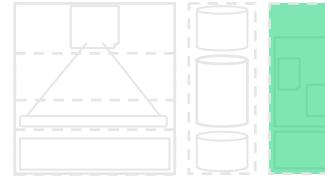
Norm **FORBIDDEN**(allocator DO *assign(organ, recipient)*)
condition **IF NOT**(hospital DONE *ensure-quality(organ)*)

Violation **NOT**(done(*ensure-quality(organ)*)) **AND**
condition *done(assign(organ, recipient))*

Detection {*detect-alarm(assign, 'starting')*;
mechanism *check(done(ensure-quality(organ)))*};

Sanction *inform(board, "NOT(done(ensure-quality(organ))*
 AND *done(assign(organ, recipient))")*

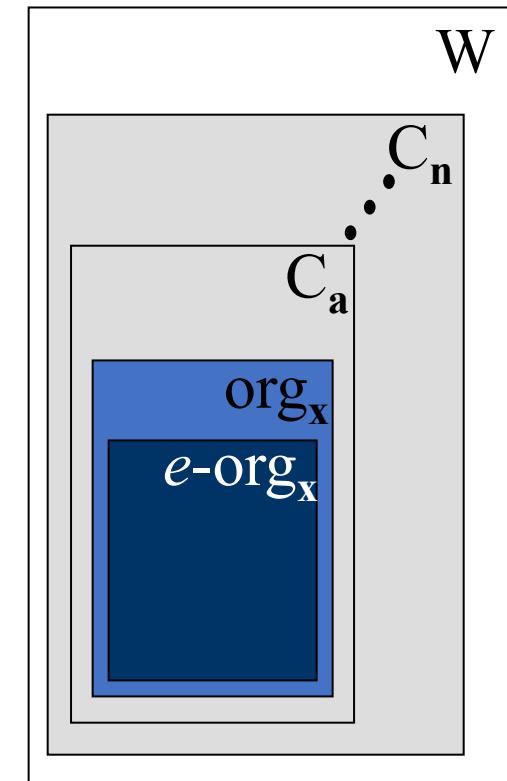
Repairs {*stop-assignation(organ)*;
 record("NOT(done(ensure-quality(organ)) AND
 done(assign(organ, recipient))", incident-log);
 detect-alarm(ensure-quality, 'done');
 check(done(ensure-quality(organ)));
 resume-assignation(organ)};



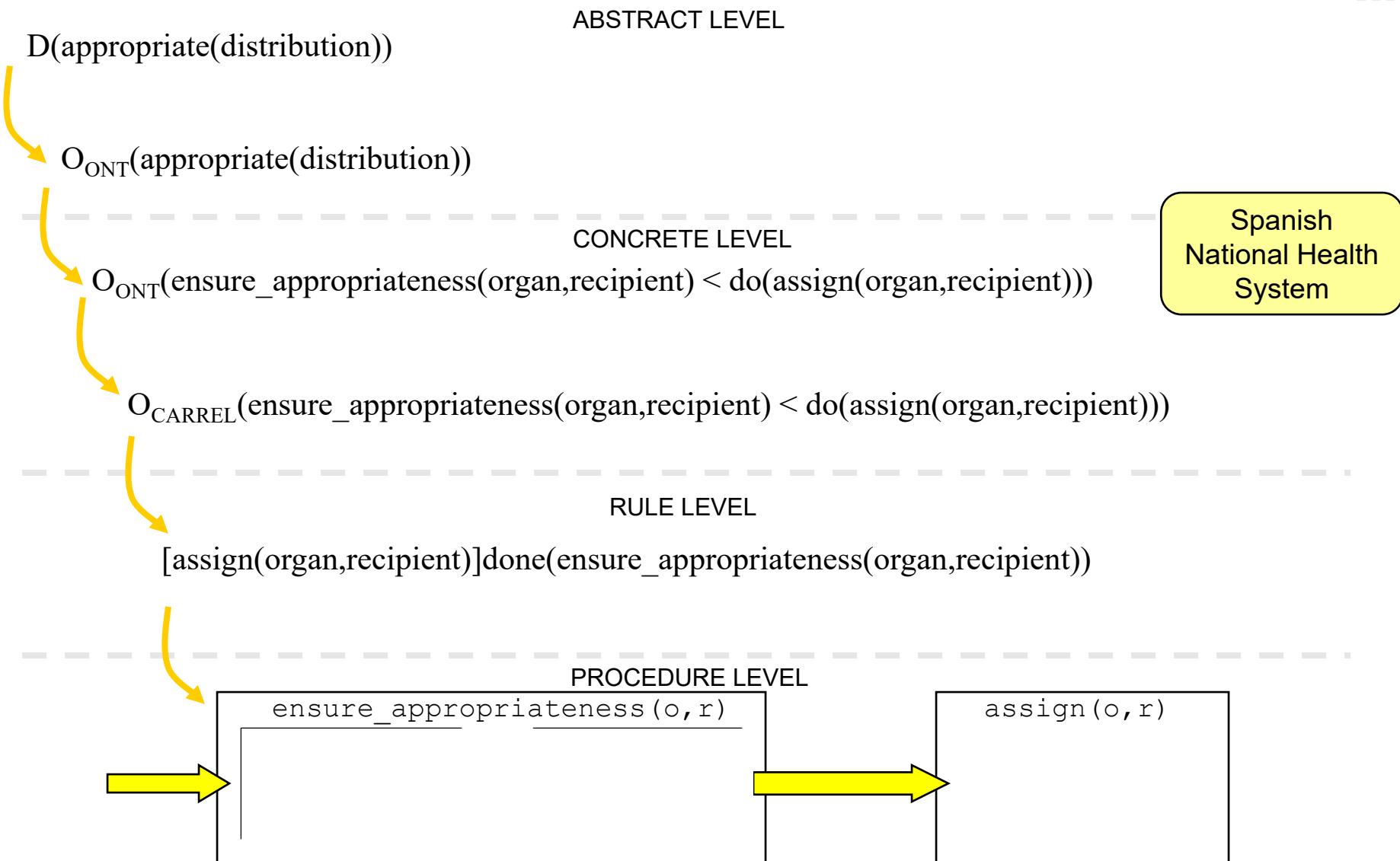
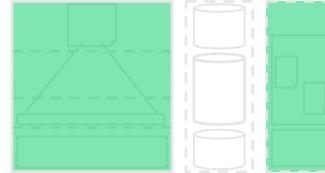
Entornos Normativos: e-Instituciones

HARMONIA – influencia del Contexto

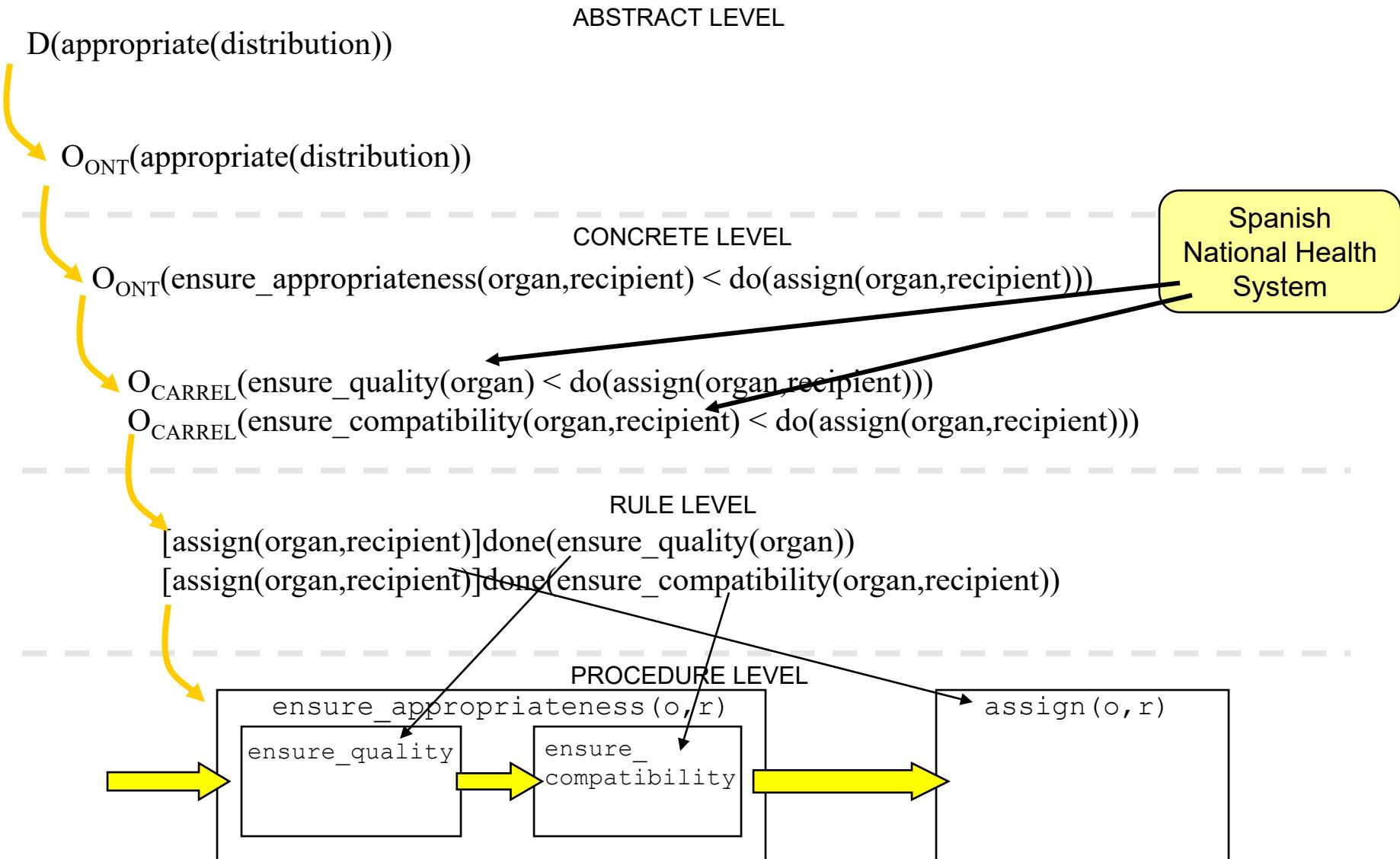
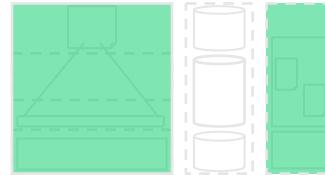
- Los **Estatutos** hacen referencia a un contexto circundante
- Vínculos con la idea de **contextos anidados**
 - $e\text{-org}_x$ es un contexto que define un **vocabulario** y un **sistema normativo**
 - Hay supercontextos que tienen una **influencia** en la definición de $e\text{-org}_x$
- influencia en varios niveles de abstracción
 - Vocabulario (términos, predicados)
 - Valores, normas, reglas y procedimientos



HARMONIA – influencia del Contexto



HARMONIA – influencia del Contexto



Combinando modelos Organizacionales y Normativos

Modelos simbólicos de IA Social

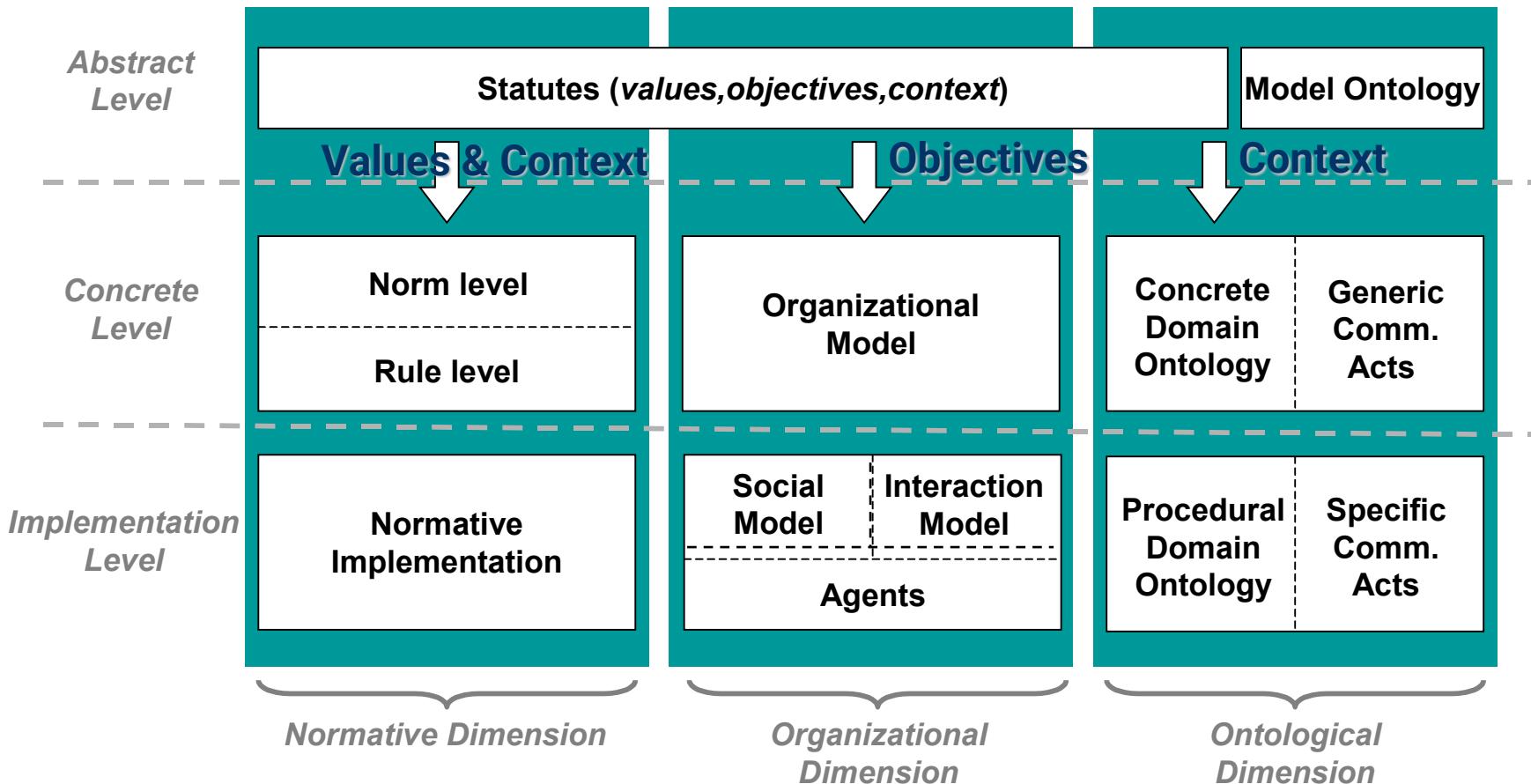
OMNI

J. Vázquez-Salceda, V. Dignum and F. Dignum

- OMNI: Organizational Model for Normative Institutions
 - Integración de los conceptos normativos de **HARMONIA** y los conceptos organizacionales de **OperA**
 - Sistemas cerrados con participantes fijos y protocolos de interacción,
 - Sistemas abiertos y flexibles que permiten y se adaptan a la participación de agentes heterogéneos con agendas diferentes.
 - Modelo multinivel
 - Abstract Level: Análisis de requisitos
 - Concrete Level: Proceso de análisis y diseño
 - Implementation Level: Especificación de diseño en una arquitectura multiagente determinada
 - Top-down
 - Orientación metodológica para el diseño
 - Bottom-up
 - Rastro de origen y motivación de las normas y protocolos

OMNI

Niveles y dimensiones



OMNI

Dimensiones

- **Normative Dimension**

- Normas y reglas que guían el comportamiento de los agentes
- Incluye un modelo de la normativa ambiental
- Proviene de **HARMONIA**

- **Organizational Dimension**

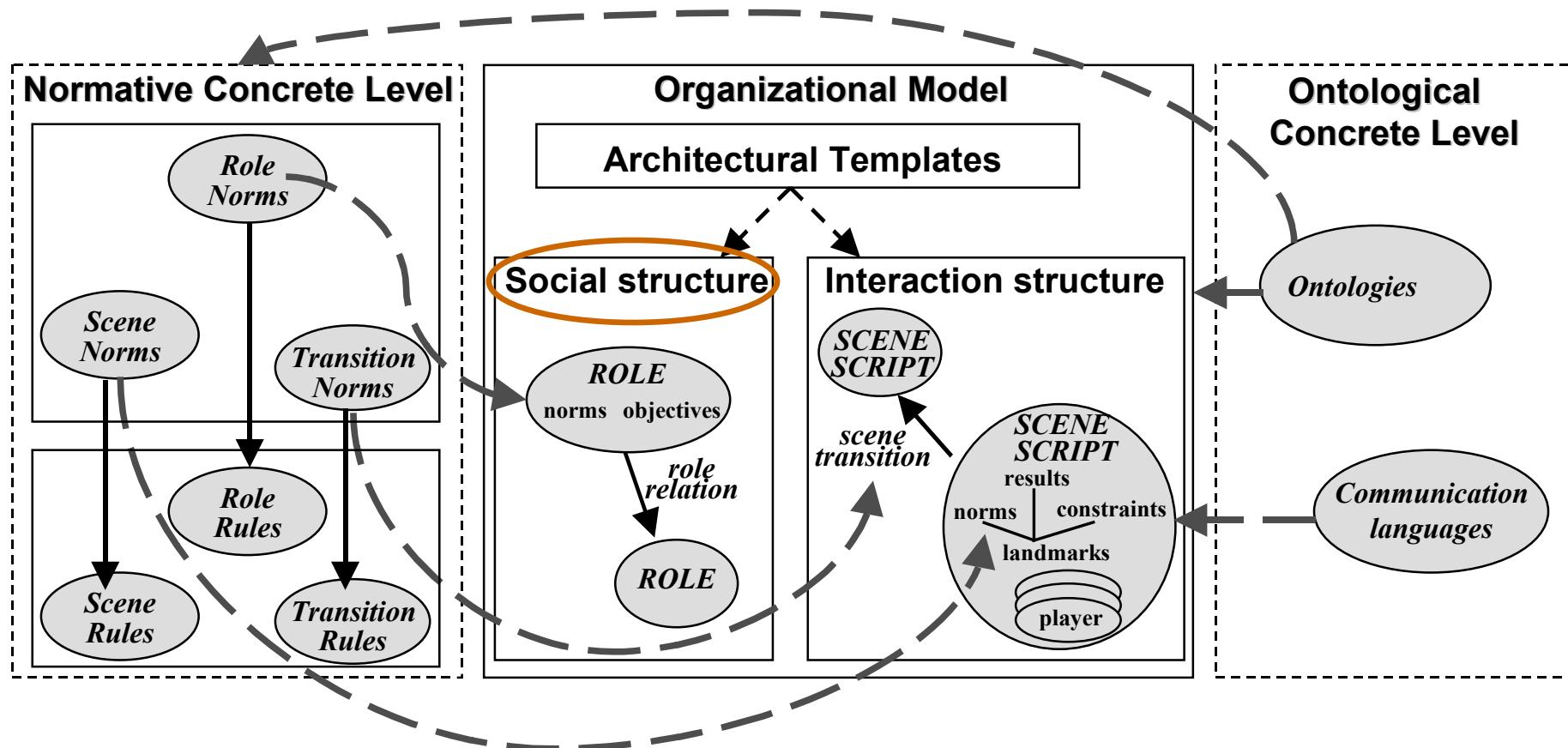
- Captura la estructura organizativa y los requisitos
- Proviene de **OperA**, con el lenguaje de normas de **HARMONIA**

- **Ontological Dimension**

- Contenido: conceptos y relaciones
- Lenguaje de comunicación
- Proviene de **OperA**

OMNI

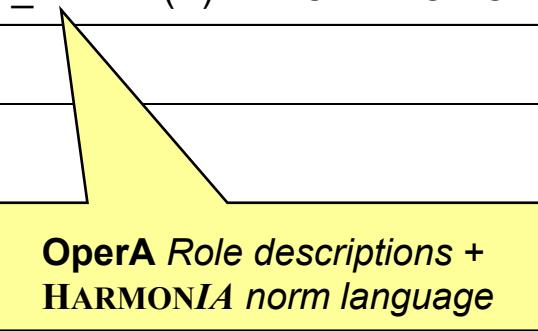
Concrete Level



OMNI

Concrete Level: Social Structure. Role example

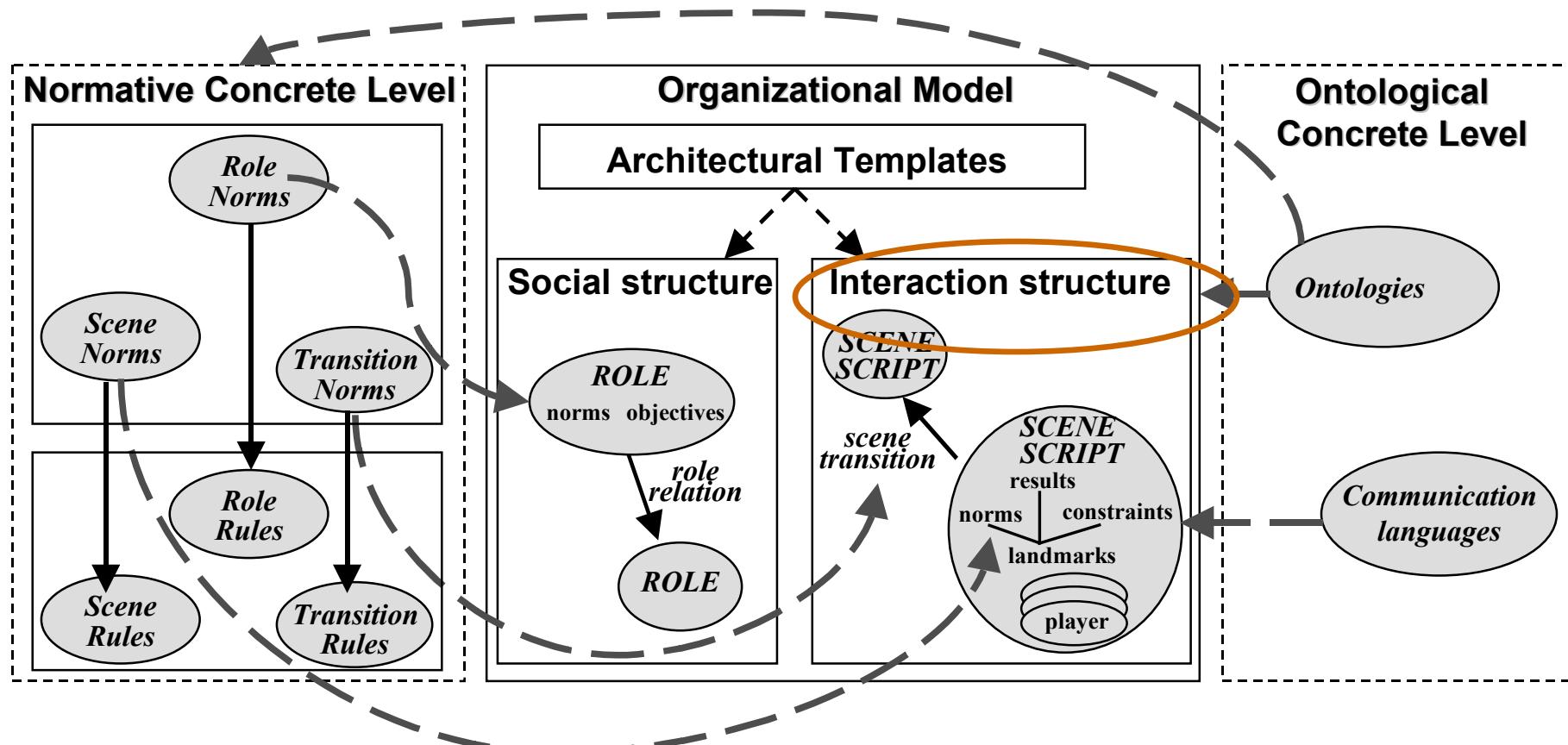
Role: PC Member	
Objectives	paper_reviewed(P, Rep)
Sub-objectives	{ read(P), report_written(P, Rep), review_received(Org, P, Rep) }
Rights	access-confman-program(me)
Norms	OBLIGED understand(English) IF DONE assigned (P, me, Deadline) THEN OBLIGED paper_reviewed(P, Rep) BEFORE Deadline IF DONE paper_assigned(P, me, _) AND direct_colleague(author(P)) THEN OBLIGED review_refused(P) BEFORE TOMORROW
Type	external



OperA Role descriptions +
HARMONIA norm language

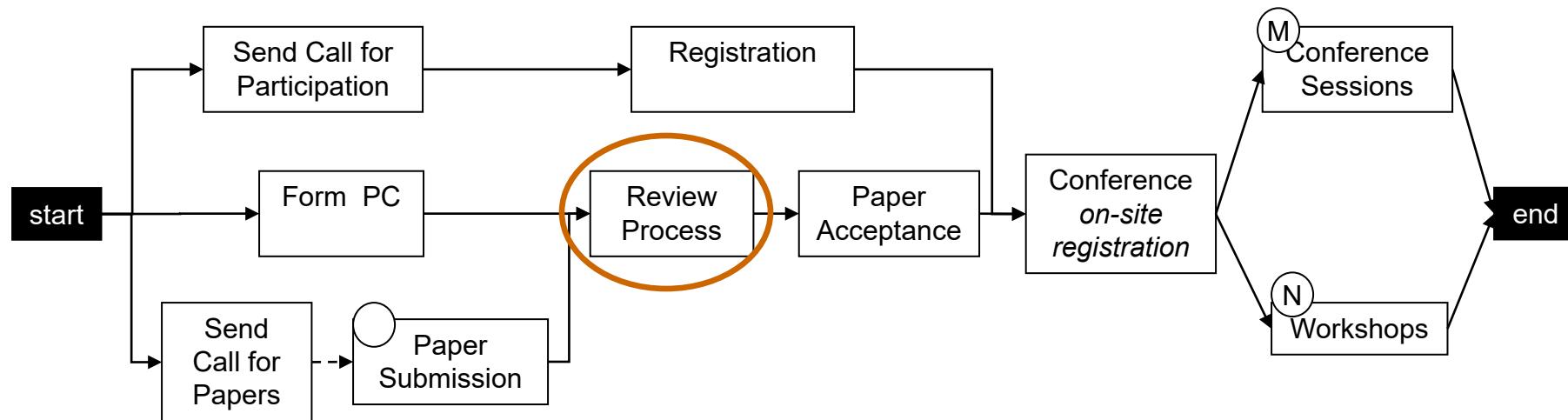
OMNI

Concrete level



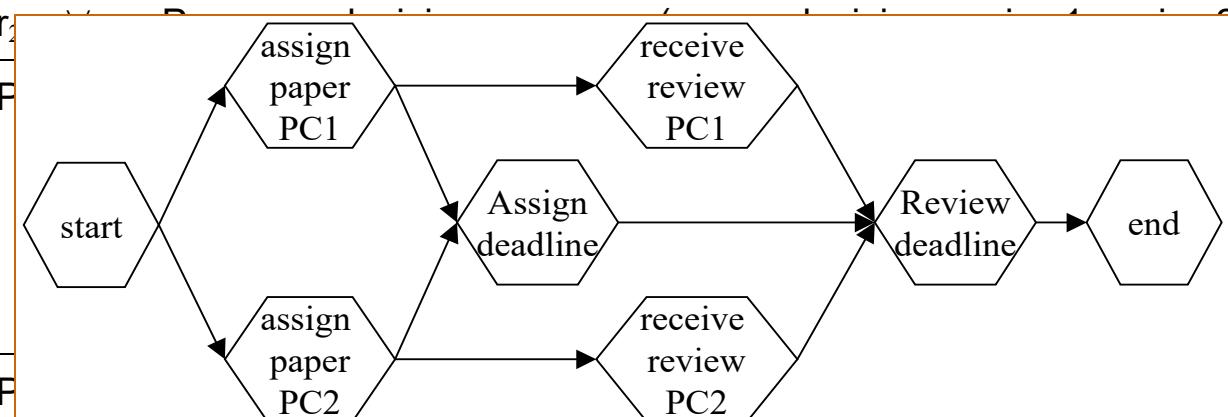
OMNI

Concrete Level: Interaction structure



OMNI

Concrete Level: Interaction Structure. Scene example

Interaction Scene: Review Process	
Roles	PC-Chair (1), PC-member (2..Max)
Results	$r_1 = \forall P \in \text{Papers}, \text{reviews_done}(P, \text{review1}, \text{review2})$ $r_2 = \forall P \in \text{Papers}, \text{reviewed}(P, \text{Rev}) \text{ BEFORE } \text{DeadlineR}$
Interaction Patterns	 <pre> graph LR start((start)) --> assign1[assign paper PC1] assign1 --> receive1[receive review PC1] receive1 --> assignD1[Assign deadline] assignD1 --> reviewD1[Review deadline] reviewD1 --> end((end)) start --> assign2[assign paper PC2] assign2 --> receive2[receive review PC2] receive2 --> assignD2[Assign deadline] assignD2 --> reviewD2[Review deadline] reviewD2 --> end </pre>
Norms	$\text{OBLIGED}(\text{PC}, \text{paper_reviewed}(P, \text{Rev}) \text{ BEFORE } \text{DeadlineR})$ $\text{OBLIGED}(\text{O}, \text{decision_on_paper}(P, D, \text{Rev1}, \text{Rev2}) \text{ BEFORE } \text{DeadlineD})$

landmarks

OperA Scene descriptions +
HARMONIA norm language

OMNI

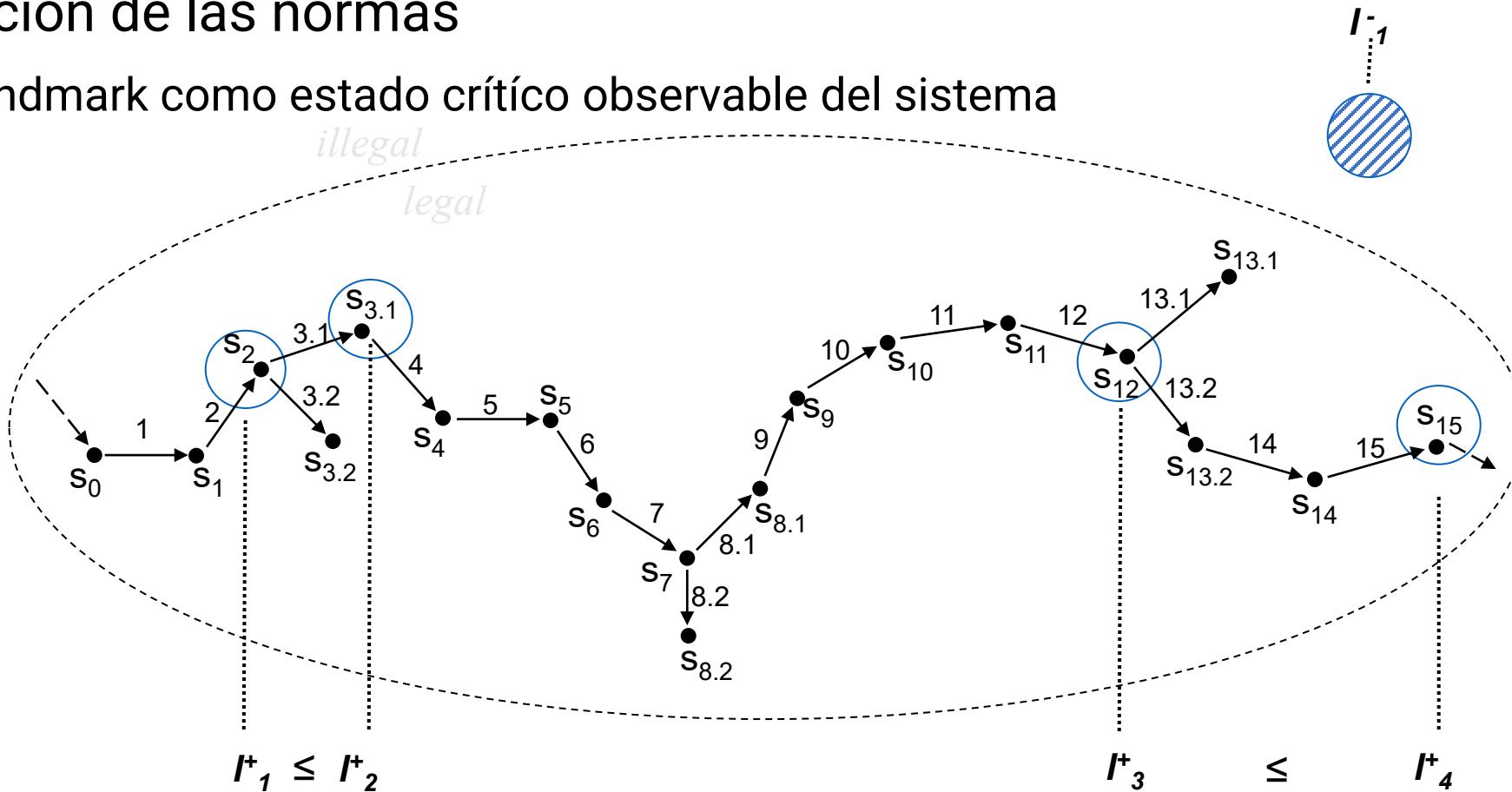
Implementation Level

- Observancia de normas
 - *Protocolos* y *Reglas* : permiten que agentes simples cumplan con las normas de la organización
- Desempeño de roles
 - *Social Contracts* : Compromisos que regulan el desempeño de funciones asociadas al rol por parte de los agentes individuales.
 - *Interaction Contracts* : Interacciones específicas, como las acordadas por los agentes.
- Ontologías
 - *Specific Communication Acts* : lenguajes de comunicación reales tal y como se fijan en los contratos de interacción.
 - *Specific Communicative Acts* : Implementar las ontologías de contenido.

OMNI

Implementation Level: Observancia de Normas

- Basado en **HARMONIA**
- Nueva idea: los Landmarks de **OperA** para monitorizar la ejecución de las normas
 - Landmark como estado crítico observable del sistema



Temas de investigación

Modelos simbólicos de IA Social

Temas de investigación

Modelos sociales computacionales enriquecidos (valores, convenciones, normas)

- Valores y Normas en el proceso de razonamiento de agentes
 - Aplicables a asistentes virtuales, Internet de la Cosas, Ambient Intelligence, Robots Sociales...
 - Puede ser la base para asegurar que sistemas de IA estén alineados con los valores éticos y el reglamento de la sociedad (por ejemplo: la IA Act Europea)
- Plataformas de agentes para ejecutar simulaciones sociales enriquecidas (con valores, convenciones, normas)
 - Evaluación “in silico” de nuevas políticas o medidas sociales (como se hizo durante el COVID-19)
 - Posiblemente ejecutadas en arquitecturas de supercomputación

Temas de investigación

Integrar/conectar modelos numéricos y simbólicos

- Hacer que agentes “de caja negra” sean explicables en entornos complejos
 - ¿Está el comportamiento de este agente RL/sistema MARL alineado con las normas y valores de la sociedad?
- Hacer que los humanos sean explicables
 - Intentar explicar el comportamiento observado de un humano (ej; en un video)
 - Aplicar modelos de explicabilidad en la interacción hombre-máquina (Human Computer Interaction)
- Aplicación de modelos sociales a métodos de explicabilidad
- Adición de agentes opacos explicables a simulaciones sociales