

38 JAIIO
38^a Jornadas Argentinas de Informática
MAR DEL PLATA, ARGENTINA
24 al 28 de agosto de 2009

Anales 38 JAIIO - ISSN 1850-2776

ASAI 2009
ISSN 1850-2784

AST 2009
ISSN 1850-2806

HPC 2009
ISSN 1851-9326

JSL 2009
ISSN 1850-2857

SID 2009
ISSN 1850-2814

SIE 2009
ISSN 1851-2526

SIS 2009
ISSN 1850-2822

SSI 2009
ISSN 1850-2830

SIO 2009
ISSN 1850-2865

EST 2009
ISSN 1850-2946

WSegI 2009
ISSN 978-987-1312-22-1

ASSE 2009
ISSN 1850-2792

CAI 2009
ISSN 1852-4850

JII 2009
ISSN 1850-2849

JUI 2009
ISSN 1851-2518

ASSE 2009
ISSN 1850-2792

CAI 2009
ISSN 1852-4850

JII 2009
ISSN 1850-2849

JUI 2009
ISSN 1851-2518

ASSE 2009
ISSN 1850-2792

CAI 2009
ISSN 1852-4850

JII 2009
ISSN 1850-2849

JUI 2009
ISSN 1851-2518



Sociedad Argentina de Informática
Buenos Aires, Argentina. Tel. (011) 4373-5733 / 4372-1950. www.sadinformatica.org.ar

JAIIO 2009 - Windows Internet Explorer

D:\html\index.html

Archivo Edición Ver Favoritos Herramientas Ayuda

Buscar web... Favoritos Spaces

JAIIO 2009

Para ayudar a proteger su seguridad, Internet Explorer impidió que esta página web ejecutara scripts o controles ActiveX que podrían obtener acceso al equipo. Haga clic aquí para ver opciones...



38 JAIIO

MAR DEL PLATA, ARGENTINA
24 AL 28 DE AGOSTO DE 2009

Inicio

Organización

Sponsors

Auspiciantes

Simposios

- ASAI 2009
- ASSE 2009
- AST 2009
- CAI 2009
- HPC 2009
- JII 2009
- JSL 2009
- JUI 2009
- SID 2009
- SIE 2009
- SIO 2009
- SIS 2009
- SSI 2009
- WSegl 2009
- EST 2009

Buscar

JAIIO 2009

JORNADAS ARGENTINAS DE INFORMÁTICA 2009 : ISSN 1850-2776

Desde 1961 se realizan las JAIIOs, Jornadas Argentinas de Informática, organizada por la SADIO, donde en sesiones paralelas se presentan trabajos que se publican en Anales, se discuten resultados de investigaciones y actividades sobre diferentes tópicos, desarrollándose también conferencias y reuniones con la asistencia de profesionales argentinos y extranjeros.

Las JAIIOs se organizan como un conjunto de simposios separados, cada uno dedicado a un tema específico, de uno o dos días de duración, de tal forma de permitir la interacción de sus participantes. Este año las 38 JAIIO se realizan en la ciudad de Mar del Plata, del 24 al 28 de agosto de 2009, en el Hotel 13 de Julio (9 de Julio 2777)

Mi equipo

Entendibilidad de Diagramas de Secuencia: Comprensión Semántica, Retención y Transferencia

Marcela Genero¹, José A. Cruz-Lemus¹, Danilo Caivano², Silvia Abrahão³, Emilio Insfran³, José A. Carsi³

¹ Grupo ALARCOS, Departamento de Tecnologías y Sistemas de Información– Universidad de Castilla-La Mancha (UCLM)
Paseo de La Universidad, 4 – 13071 – Ciudad Real – España

²Departamento de Informática – Universidad de Bari
Via E. Orabona 4 – 70126 – Bari – Italia

³ Grupo ISSI, Departamento de Sistemas Informáticos y Computación
Universidad Politécnica de Valencia (UPV) – Valencia – España

{Marcela.Genero,JoseAntonio.Cruz}@uclm.es, caivano@di.uniba.it, {sabrahao, einsfran, pcarsi}@dsic.upv.es

Resumen. En UML, los estereotipos se utilizan para precisar la semántica requerida para un elemento ya existente, o para definir un nuevo elemento gráfico con su respectiva semántica. Sin embargo, las posibles mejoras introducidas por la utilización de dichos mecanismos de extensión en los diagramas de secuencia aún no cuentan con un consenso generalizado. En este artículo analizamos la efectividad del uso de estereotipos como mecanismo de mejora de la entendibilidad de diagramas de secuencia UML mediante tres medidas: comprensión semántica, retención y transferencia. Los resultados obtenidos indican que, en general, los diagramas estereotipados pueden ser más comprensibles, menos propensos a errores y requerir menos tiempo para ser entendidos que los diagramas que no los utilizan.

Palabras claves: UML, diagramas de secuencia, estereotipos, entendibilidad, experimento

1. Introducción

La aparición del paradigma de desarrollo de software dirigido por modelos ha hecho que el foco de interés de la calidad del software, se haya desplazado desde el código a los modelos conceptuales. Específicamente, a los modelos construidos utilizando el estándar de modelado “Lenguaje Unificado de Modelado” (UML) [1]. Sin embargo, el principal inconveniente de UML es que el conjunto de constructores de modelado que utiliza no está pensado para un propósito específico y por ello el modelado puede no ser efectivo [2]. Para evitar este inconveniente, varios trabajos proponen mejorar UML mediante mecanismos de extensión del lenguaje (estereotipos).

Los estereotipos normalmente se utilizan para aclarar o extender el significado de los elementos del modelo. Sin embargo, en los últimos años se han realizado varios estudios en el campo de la Ingeniería del Software para comprobar la efectividad del uso de los estereotipos

[2, 3]. Estos estudios investigan la influencia del uso de estereotipos en la entendibilidad de los diagramas de clases y colaboración UML. La entendibilidad es uno de los principales factores que afectan a la mantenibilidad [4, 5, 6, 7] y debe valorarse desde las primeras etapas del ciclo de vida del software. Sin embargo, la influencia de los estereotipos en la entendibilidad de especificaciones de requisitos no ha sido investigada aún en profundidad.

En este artículo, analizamos los resultados de un experimento controlado para evaluar si el uso de estereotipos mejora la entendibilidad de los diagramas de secuencia UML. Concretamente, se considera un conjunto de estereotipos que han sido propuestos para enriquecer la semántica de los mensajes de interacción en los diagramas de secuencia UML [8, 9], en el contexto de un método de Ingeniería de Requisitos para el desarrollo software dirigido por modelos. Resultados preliminares de este trabajo, correspondiente a un estudio piloto para validar los materiales experimentales, han sido presentados en [10].

El propósito de la utilización de los estereotipos en los diagramas de secuencia es doble. En primer lugar, pretenden mejorar la entendibilidad de los diagramas de secuencia UML, dotando a los mensajes de una semántica más clara y facilitar, así, la comunicación entre los desarrolladores y los *stakeholders*. En segundo lugar, facilitan la transformación de los diagramas de secuencia en modelos conceptuales en el marco de desarrollo dirigido por modelos, mejorando la trazabilidad entre el modelo de requisitos (basado en diagramas de secuencia) y los modelos de análisis y diseño.

La estructura del artículo es la siguiente. La sección 2 describe el estado del arte en cuanto a la evaluación empírica de la entendibilidad de diagramas UML respecto al uso de estereotipos. La sección 3 presenta la definición de los estereotipos para diagramas de secuencia UML. La sección 4 describe la planificación y ejecución del experimento. La sección 5 presenta el análisis de datos y la interpretación de los resultados. Finalmente, la sección 6 describe las conclusiones y trabajos futuros.

2. Trabajos Relacionados

La entendibilidad de modelos UML ha sido frecuentemente investigada. Entre los trabajos más representativos se encuentran los siguientes estudios: [4, 5, 11, 12, 13, 14, 15, 16, 17].

En [14, 15] Purchase et al. describen dos experimentos que abordan el entendimiento de los diagramas de clases y colaboración de UML. En el caso de los diagramas de clases, obtuvieron que el rendimiento de las notaciones depende de las tareas para las cuales se usan los diagramas. En un trabajo similar, Otero y Dolado [5] evalúan la entendibilidad del modelado dinámico en el diseño basado en UML usando dos experimentos en los que comparan el entendimiento de los diagramas de secuencia, colaboración y estados UML.

Cruz-Lemus et al. [11] presenta una familia de experimentos para investigar si el uso de los estados compuestos mejora la entendibilidad de los diagramas de estados UML. Los resultados muestran que el uso de estados compuestos no mejoran la entendibilidad de estos diagramas. Glezer et al. [12] compara experimentalmente la entendibilidad de dos diferentes diagramas de interacción UML (colaboración y secuencia) en dos dominios de aplicación, los sistemas de gestión de la información (MIS) y sistemas de tiempo real (RTS). Los resultados muestran que, en general, los diagramas desarrollados para sistemas MIS parecen ser más fácil de comprender que los diagramas para sistemas RTS.

Genero et al. [4] presentan una familia de experimentos para evaluar la relación de un conjunto de medidas para diagramas de clases UML y su entendibilidad y capacidad de

modificación. Los resultados muestran que las medidas relacionadas con asociaciones y generalizaciones tienen una mayor influencia en la entendibilidad y modificación de los diagramas. Lange y Chaudron [13] validan a través de un experimento si existe una diferencia entre un conjunto de vistas que ellos proponen para incrementar la comprensión de modelos y las vistas existentes con respecto a la correcta comprensión y esfuerzo. Concluyen que la corrección en la entendibilidad mejora y el tiempo necesario para hacerlo se reduce al usar las vistas que ellos proponen.

Xie et al. [16] presentan un estudio en el que proponen una notación para diagramas de secuencia UML que incluye sincronización y la usan para comprobar si ayuda a los alumnos a entender mejor los conceptos de ejecución simultánea y concurrencia. Finalmente encuentran un beneficio significativo. Por último, Yusuf et al. [17] identifican las características específicas de los diagramas de clases UML que son más efectivas para apoyar una tarea dada (por ejemplo, presentación, color, uso de estereotipos, etc.). Los resultados indican que la realización de las tareas depende de la habilidad de los sujetos para diseñar software.

Con respecto al uso de estereotipos para mejorar la entendibilidad de modelos UML, que nosotros sepamos, existen dos trabajos recientes [2, 3]. Staron et al. [2] valoran la influencia de estereotipos en la entendibilidad de modelos (diagramas de clases y colaboración UML) a través de un conjunto de experimentos. Encuentran que los estereotipos ayudan significativamente tanto a alumnos como a programadores profesionales. Ricca et al. [3] describen una familia de tres experimentos realizados para evaluar la efectividad de los estereotipos UML para apoyar la comprensión de tareas en los diagramas de clases para el diseño Web. Concluyen que los estereotipos mejoran significativamente la realización en sujetos con poca experiencia.

La revisión bibliográfica descrita aquí revela que la entendibilidad es un tema de interés en el contexto del modelado UML, pero la contribución de los estereotipos en la mejora de la entendibilidad de los diagramas de secuencia UML aún no ha sido investigada en profundidad.

3. Un Conjunto de Estereotipos para Diagramas de Secuencia

Los diagramas de secuencia se usan para modelar un aspecto del comportamiento dinámico del sistema [18]. Pueden usarse en el contexto de un sistema completo, de un subsistema o de un caso de uso. Varios autores indican que la funcionalidad de los casos de uso puede ser descrita por uno o varios diagramas de secuencia [18, 19]. Cuando un diagrama de secuencia se utiliza para especificar el comportamiento de un caso de uso, la descripción del caso de uso puede utilizarse para guiar la descripción del contenido del diagrama de secuencia. Durante todo el proceso de análisis de requisitos, el diagrama de casos de uso puede ser revisado en base a los resultados del diagrama de secuencia y viceversa, hasta que ambos estén sincronizados apropiadamente [20].

Basándonos en la clasificación de escenarios propuesta en [21], cuando se usa un diagrama de secuencia a nivel de tipo, las entidades representadas no son objetos sino tipos de objetos (clases¹) y representan un modelo de interacción, es decir, un conjunto de mensajes entre los tipos de objetos para realizar un comportamiento. Un diagrama de secuencia cuenta con dos dimensiones: la dimensión vertical, que representa el tiempo, y la dimensión horizontal, que representa los diferentes tipos de objetos que participan.

¹ En este trabajo, los términos *clase* y *tipo de objeto* se utilizan indistintamente.

El énfasis en este tipo de diagramas consiste en poder representar gráficamente el modelo de interacción entre los tipos de objetos enviando y recibiendo mensajes a medida que el tiempo avanza. Además, cuando usamos diagramas de secuencia para realizar un caso de uso, es decir, a nivel de análisis de requisitos, el objetivo no debe ser la especificación en detalle del comportamiento con iteraciones complejas y mensajes condicionales (el cómo) sino las iteraciones que sean necesarias para realizar el propósito del caso de uso (el qué).

Para construir el diagrama de secuencia de un caso de uso, se debe analizar el caso de uso a dos niveles: nivel de diagrama de caso de uso (los actores que se comunican con el caso de uso) y nivel de especificación del caso de uso (el conjunto de pasos o responsabilidades que deben hacerse a fin de llevar a cabo el caso de uso).

A nivel de diagrama de caso de uso, el actor es un elemento externo que puede enviar y/o recibir información del caso de uso. Cada caso de uso tiene al menos un actor, que inicia el caso de uso y otros actores colaboradores, que también pueden enviar y/o recibir información.

A nivel de especificación del caso de uso, la principal tarea es obtener el conjunto de interacciones necesarias para realizar el propósito del caso de uso (descrito textualmente en forma de pasos o responsabilidades del caso de uso). Esencialmente, cuando el caso de uso recibe un estímulo (una interacción externa), el sistema produce un conjunto de interacciones entre sus componentes internos (objetos) como respuesta. Estas interacciones se representan gráficamente mediante mensajes entre tipos de objetos en el diagrama de secuencia.

En [8, 9] se presenta una clasificación específica para estos mensajes, teniendo en cuenta la naturaleza de los mismos y se identifican usando los siguientes estereotipos: «signal», «service», «query» y «connect».

Los mensajes de tipo «signal» representan interacciones entre un actor (un tipo de entidad externa) y el sistema. Los actores se comunican con el sistema enviando estímulos al interfaz del sistema, y el interfaz envía respuestas a los actores. Los mensajes de tipo «service» son mensajes que cambian el estado de los objetos que componen el estado del sistema. En este tipo de mensajes, el origen del mensaje puede ser una clase del sistema o el interfaz y el destino del mensaje es siempre una clase del sistema. Los mensajes de tipo «query» representan consultas sobre otros objetos o sobre la población de una clase. En el contexto de los diagramas de secuencia, este tipo de interacción es el único mecanismo usado para conocer el estado de otro objeto. Por último, los mensajes de tipo «connect» capturan un importante tipo de interacción entre los objetos participantes (aunque a veces es difícil identificar interacciones de este tipo a primera vista). En el modelo orientado a objetos, un objeto puede estar formado por otros objetos o puede estar asociado (relacionado) con otros objetos. Debido al concepto de encapsulación, estos dos casos son el único camino para que un objeto pueda tener acceso al estado de otro objeto. Por esta razón, cuando se crea el objeto (o más tarde en la vida del objeto), se debe indicar todos sus otros objetos relacionados, y por tanto se produce una interacción entre éstos. Por este motivo, y dado que usamos diagramas de interacción, específicamente diagramas de secuencia, las interacciones de este tipo deben ser tenidas en cuenta y representadas adecuadamente (en este caso, usando el estereotipo «connect»).

Con respecto al uso de los estereotipos, en el estándar UML 2.1 [1] se define un conjunto general de estereotipos predefinidos que pueden ser usados en modelos UML dependiendo del nivel de conformidad con el estándar (L2 o L3), pero en este conjunto no existen estereotipos para los diagramas de secuencia. Por esta razón, métodos y herramientas específicos deben extender y adaptar la notación y el significado de los elementos del estándar UML para tener un modelo con mayor riqueza semántica y de notación.

En la Figura 1, se presenta un ejemplo que muestra las interacciones estereotipadas (mensajes) que son necesarias para realizar un caso de uso de un sistema de venta de artículos en una tienda convencional. La explicación del significado de cada una de las interacciones se proporciona a continuación:

- El cajero comienza la venta. El mensaje 1 representa el comienzo de la comunicación entre el cajero y el sistema a través de su interfaz. El mensaje 2 (*introducir_datos_venta*) proporciona la información necesaria para realizar la operación.
- El cajero crea una nueva venta (mensaje 3). Este mensaje implica la responsabilidad de crear una nueva venta y registrar su información. Se llama a esta interacción *crear_venta* y el destinatario es un nuevo objeto de la clase *Venta*.
- Se registra la información sobre cada artículo que se vende (mensaje 4). Este mensaje implica la responsabilidad de registrar la información del artículo a vender. Se llama a esta interacción *crear_linea_de_venta* y el destinatario es un nuevo objeto de la clase *LíneaDeVenta*.
- Se identifica el artículo a ser vendido (mensaje 5). Este mensaje implica la responsabilidad de que la *LíneaDeVenta* debe tener uno y sólo un *Artículo* asignado. Se llama a esta interacción *seleccionar_artículo* y el destinatario es un objeto existente de la clase *Artículo*.

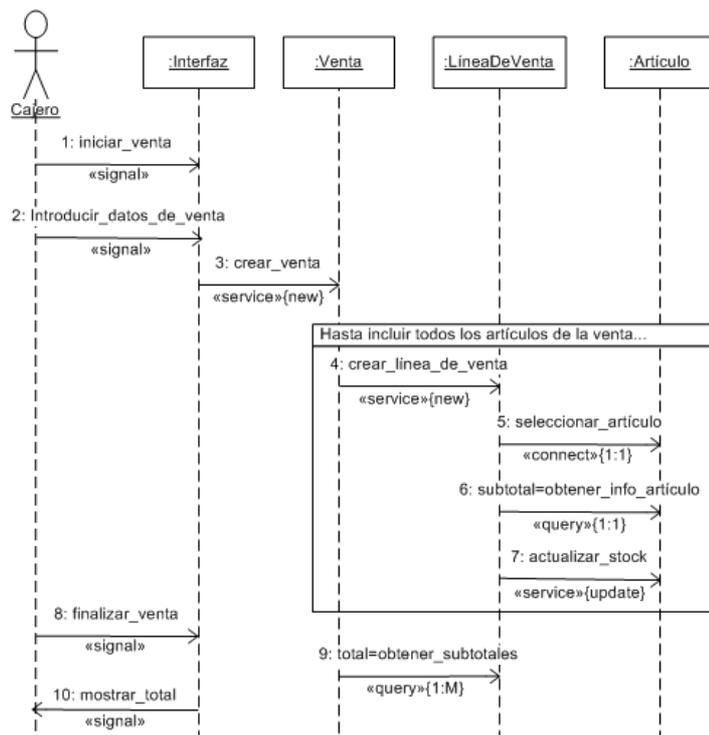


Figura 1. Ejemplo de diagrama de secuencia con estereotipos

- Se determina el precio del *Artículo* vendido y añadir esta información a la *LíneaDeVenta* actual (mensaje 6). Este mensaje implica la responsabilidad de preguntar a un objeto de la clase *Artículo* su precio almacenado. Se llama a esta interacción *obtener_info_articulo*.
- Se actualiza el stock del *Artículo* vendido (mensaje 7). Este mensaje implica la responsabilidad de modificar el estado del objeto *Artículo* vendido. Se llama a esta interacción *actualizar_stock*.
- El cajero indica que la venta ha acabado enviando un mensaje del tipo «signal» al interfaz del sistema (mensaje 8).
- Se calcula el precio total de la venta consultando el conjunto de artículos vendidos (objetos de la clase *LíneaDeVenta*). Esto implica la responsabilidad de que la venta interactúe con su conjunto de líneas de venta para obtener los subtotales correspondientes (mensaje 9). Finalmente, la siguiente interacción muestra la información de la venta realizada al cajero (mensaje 10).

Cabe recordar que a este nivel de análisis de requisitos, los detalles de la iteración o del control de flujo en el diagrama de secuencia son irrelevantes ya que sólo se está especificando qué interacciones son necesarias para realizar el caso de uso. Este tipo de actividad nos dirige hacia el descubrimiento de todas las clases del sistema y sus interrelaciones, a mejorar el entendimiento del problema restringido a un caso de uso, y no a describir el comportamiento detallado de las clases involucradas.

De esta forma, el proceso de especificación de un caso de uso continúa hasta que todos los pasos definidos en la especificación textual del caso de uso tienen su correspondencia como interacciones en el diagrama de secuencia. Este proceso pretende que el caso de uso pueda evaluarse junto con los usuarios durante y después de su construcción. Por tanto, la entendibilidad es un factor crítico en la especificación de este tipo de diagramas.

4. Diseño del Experimento

El objetivo de este experimento es evaluar si el uso de estereotipos mejora la entendibilidad de los diagramas de secuencia. Por eso, usando la plantilla GQM para la definición de objetivos [22, 23], el objetivo del experimento se define como: *analizar* el uso de estereotipos *con el fin de evaluarlos con respecto a la* entendibilidad de los diagramas de secuencia UML *desde el punto de vista de* los investigadores, *en el contexto de* alumnos de cuarto curso de informática en la Universidad de Bari, en Italia.

Para realizar este experimento se siguen las recomendaciones proporcionadas en varios trabajos [24, 25]. El diseño del experimento presentado en este artículo es similar al presentado en [2]. Seleccionamos un diseño factorial balanceado con interacción confundida en los grupos [26]. Los objetos fueron diagramas de secuencia UML, con dos posibles valores: diagramas con estereotipos y diagramas sin estereotipos.

Todos los alumnos fueron asignados a 4 grupos aleatoriamente (grupo 1, 2, 3 y 4). Hubo dos rondas en el experimento. Dos diagramas diferentes se presentaron a cada alumno en cada grupo. Para evitar un posible efecto de aprendizaje, los diagramas correspondían a diferentes dominios de aplicación (A y B). Los diagramas eran similares en complejidad y los dominios eran conocidos por los alumnos.

4.1. Planificación

El experimento fue llevado a cabo por setenta y siete alumnos de cuarto curso de Informática que habían realizado un curso de Ingeniería del Software Empírica. Los alumnos habían cursado dos asignaturas de Ingeniería del Software en las cuales adquirieron habilidades para trabajar con diagramas UML. El conocimiento que tenían era el necesario para entender los diagramas sin estereotipos, y todos ellos tenían la misma formación. Conocían el uso de los estereotipos en general, pero se les enseñaron los estereotipos propuestos para diagramas de secuencia UML en una sesión de entrenamiento organizada un día antes de la ejecución del experimento.

Para evitar rechazos debido a la aprensión que pudiera producirles una evaluación, los alumnos no fueron calificados según su actuación. A los participantes se les concedieron dos puntos extra en la evaluación final del curso. Los objetos experimentales consistieron en cuatro diagramas:

- A-C: diagrama A con estereotipos y una descripción de cada tipo de estereotipo.
- A-S: diagrama A sin estereotipos.
- B-C: diagrama B con estereotipos y una descripción de cada tipo de estereotipo.
- B-S: diagrama B sin estereotipos.

Cada diagrama de secuencia UML A-x describe el dominio de un alquiler de coches. En concreto, describe la especificación de un caso de uso “Alquiler de extra” que consiste en el alquiler de artículos adicionales al alquiler de un coche, por ejemplo GPS, DVD portátil, etc. Cada diagrama B-x describe el dominio de un hotel. En concreto, describe la representación de un caso de uso “Reserva de habitaciones”. Para evitar problemas de interpretación, los objetos experimentales se presentaron en italiano, lengua habitual de los sujetos. El material experimental completo puede encontrarse en <http://alarcos.esi.uclm.es/ExpStereotypes>

Existen dos variables independientes en el experimento, el tipo de diagrama, con valores: C (con estereotipos) y S (sin estereotipos), y el dominio del diagrama (A y B). Combinando cada nivel de variables independientes tenemos cuatro opciones, como se refleja en los cuatro diagramas que son los objetos del experimento.

En principio cogimos como variable dependiente a la entendibilidad de los diagramas de secuencia. Según algunas sugerencias sobre cómo medir la entendibilidad [11, 27, 28], hemos usado la Teoría Cognitiva de Aprendizaje Multimedia (CTML) [18]. Esta elección se basa en varias razones, como por ejemplo el hecho de que se centre en el uso conjunto de palabras y gráficos, que son los elementos de la gramática de los diagramas de secuencia UML, o el de ser una teoría que ha evolucionado durante años de trabajo y desarrollo de instrumentos y métodos experimentales [29, 30].

Siguiendo la CTML, la entendibilidad de los diagramas de secuencia UML se ha definido a través de tres variables:

- *Comprensión semántica*: la habilidad para comprender la semántica de los modelos.
- *Retención*: la entendibilidad del material que ha sido dado y la habilidad para retener (memorizar) conocimiento de él.
- *Transferencia*: la habilidad para usar el conocimiento adquirido del material para resolver problemas relacionados con él pero que no pueden responderse directamente con el material experimental.

Para medir estas variables, hemos usado tres tests distintos basados en cuestionarios o tareas. Cada medida de entendibilidad se calcula como:

- Efectividad: proporción de respuestas correctas proporcionadas en cada test (número de respuestas correctas / número de cuestiones). Esta medida refleja la habilidad para entender el material proporcionado correctamente.
- Eficiencia: proporción de respuestas correctas por tiempo (Efectividad / Tiempo)

En este trabajo, se llama a estos valores CSEfec/CSEfic para la efectividad y la eficiencia de la comprensión, RetenEfec/RetenEfic para la efectividad y eficiencia de la retención y, finalmente, TransEfec/TransEfic para la efectividad y la eficiencia de la transferencia. Para la comprensión semántica se comprueban las siguientes hipótesis:

- **Hipótesis nulas:**

$H_{1.1,0}$: los estereotipos no mejoran la CSEfec de los materiales cuando se trata de comprender un diagrama de secuencia UML.

$H_{1.2,0}$: los estereotipos no mejoran la CSEfic de los materiales cuando se trata de comprender un diagrama de secuencia UML.

- **Hipótesis alternativas:**

$$H_{1.1,1} = \neg H_{1.1,0}$$

$$H_{1.2,1} = \neg H_{1.2,0}$$

Análogamente, formulamos un conjunto de hipótesis H2 para las medidas de retención (RetenEfec y RetenEfic), y otro conjunto H3 relacionado con las medidas de transferencia (TransEfec y TransEfic).

La instrumentación usada en el experimento consistió en tres tests relacionados con cada uno de los cuatro tratamientos. Se indicó a los alumnos registrar el tiempo de inicio y fin de cada test. Los tests se describen a continuación:

- Test 1: contiene 10 preguntas de Sí/No sobre la semántica de los diagramas. A partir de estas preguntas fueron obtenidos la CSEfec y la CSEfic.
- Test 2: consiste en preguntas de “rellenar huecos” donde los alumnos tuvieron que rellenar un texto que describía la funcionalidad de los diagramas. A partir de estas preguntas fueron calculadas la RetenEfec y la RetenEfic.
- Test 3: los alumnos tenían que dar nombre a un conjunto de nuevos mensajes de interacción que habían sido introducidos en la versión original de los diagramas, pero etiquetados sólo con parámetros. A partir de estas preguntas fueron calculados la TransEfec y la TransEfic.

4.2. Desarrollo del Experimento

En el transcurso del experimento, hubo dos rondas. En cada ronda, a cada uno de los grupos les fue dado un tratamiento distinto. Para evitar el efecto aprendizaje, dividimos a los sujetos en cuatro grupos. Aleatoriamente asignamos los diagramas correspondientes a cada grupo, pero en diferente orden. La Tabla 1 presenta el resumen del desarrollo del experimento.

Tabla 1. Rondas del experimento

Ronda 1		Tipo de Diagrama	
		Con estereotipos	Sin estereotipos
Dominio	A	Grupo 1	Grupo 2
	B	Grupo 4	Grupo 3
Ronda 2		Tipo de Diagrama	
		Con estereotipos	Sin estereotipos
Dominio	A	Grupo 3	Grupo 4
	B	Grupo 2	Grupo 1

La descripción de cada tipo de diagrama fue presentada en la sección 4.1. Esta asignación corresponde al diseño factorial balanceado con grupo de interacción confundido [26], para evitar efectos de aprendizaje. El diseño del experimento incluye una clase previa de tres horas llevada a cabo el día anterior de la realización del experimento, que incluyó lo siguiente: (i) introducción sobre los estereotipos definidos para los diagramas de secuencia UML; (ii) introducción de un ejemplo, similar al material usado en el experimento. Se ha resuelto con los alumnos las cuestiones requeridas en el ejemplo; (iii) desarrollo de un estudio piloto para asegurarse que los alumnos han entendido el material experimental y las tareas.

Después de la sesión previa, comprobamos las respuestas y descubrimos que habían tenido una mala actuación al resolver la segunda tarea, por lo que antes de la realización del experimento, se volvió a explicar cuidadosamente las tareas del test 2.

4.3. Ejecución del Experimento

El día siguiente a la sesión de entrenamiento se realizó el experimento en una sesión de una hora y media. En los primeros 30 minutos se explicó cómo realizar el experimento y aleatoriamente se asignaron los alumnos a dos grupos balanceados. El experimento se realizó en una única aula, donde los alumnos fueron supervisados y no se permitió ninguna comunicación entre ellos.

Cada ronda (ver Tabla 1) se realizó de la siguiente forma:

- 1) Los sujetos recibieron el material de del Test1 que incluía un diagrama de secuencia y un cuestionario. Una vez que terminaban este cuestionario, los sujetos tenían que entregar el Test1 al supervisor.
- 2) Los sujetos recibieron el Test2 que debían resolver sin el diagrama de secuencia. Una vez que lo terminaban, debían entregarlo al supervisor.
- 3) Finalmente recibieron el Test3, que una vez resuelto debían entregar al supervisor.

5. Análisis e Interpretación de Datos

Se ha usado el paquete estadístico SPSS [31] para realizar los análisis estadísticos. En esta sección se presenta, para cada variable, los estadísticos descriptivos y los resultados de los tests ANOVA que se han llevado a cabo para comprobar las hipótesis formuladas. ANOVA es el test más apropiado para la exploración de los resultados de un diseño factorial con interacción confundida [26, 32].

En todos los test, usamos un umbral estadístico de significación $\alpha = 0,05$, por lo que se pueden rechazar las hipótesis nulas si los tests proporcionan un resultado estadístico significativo que no fuera mayor que 0,05.

La Tabla 2 presenta los estadísticos descriptivos para todas las variables que se han estudiado: comprensión semántica, retención y transferencia.

Tabla 2. Estadísticos descriptivos de las variables dependientes (media y desviación estándar)

Tipo	Dom	Comprensión Semántica				Retención				Transferencia			
		CSEfec		CSEfic		RetenEfec		RetenEfic		TransEfec		TransEfic	
		Med	Dev	Med	Dev	Med	Desv	Med	Dev	Med	Dev	Med	Desv
Con Est.	A	0,756	0,141	0,0020	0,000	0,636	0,157	0,0029	0,003	0,682	0,250	0,002	0,001
	B	0,782	0,131	0,0027	0,001	0,784	0,183	0,0035	0,001	0,829	0,240	0,009	0,005
	Total	0,769	0,136	0,0023	0,001	0,709	0,185	0,0032	0,002	0,755	0,254	0,005	0,005
Sin Est.	A	0,761	0,107	0,0022	0,000	0,637	0,181	0,0024	0,001	0,616	0,224	0,002	0,001
	B	0,741	0,139	0,0024	0,000	0,807	0,140	0,0036	0,004	0,821	0,313	0,009	0,006
	Total	0,751	0,124	0,0023	0,000	0,723	0,182	0,0030	0,003	0,719	0,290	0,006	0,005

Analizando la Tabla 2, se puede observar que (ver las celdas en negrita):

- **Comprensión semántica:** en ambas variables (CSEfec y CSEfic), los sujetos obtuvieron mejores resultados, en promedio, en los diagramas con estereotipos. Obviamente, esto significa que los diagramas con estereotipos resultan más entendibles, son menos propensos a errores y requieren menos tiempo para ser entendidos que los diagramas que no los utilizan.
- **Retención:** los sujetos obtuvieron mejores resultados de efectividad cuando no usaban estereotipos, mientras que las puntuaciones aumentaban en eficiencia cuando trabajaron con diagramas que incluían estereotipos. Además, los valores de RetenEfec y RetenEfic son significativamente mejores en los diagramas del dominio B. Luego, los alumnos que usaron diagramas sin estereotipos fueron capaces de recordar una mayor cantidad de funcionalidades correctas de los modelos, aunque si relacionamos el número de respuestas correctas con el tiempo invertido en contestar el test lo hicieron mejor los sujetos que utilizaron diagramas con estereotipos. Esto se debe probablemente a que los diagramas con estereotipos contienen más información que los que no los usan y por ello resultan más difíciles de memorizar (¡pero no de entender!). En este sentido, es interesante señalar que la efectividad se calcula como el número de respuestas correctas entre el número total de las respuestas esperadas y que en nuestro análisis las respuestas sin contestar son consideradas como erróneas. Se les dijo a los sujetos que no rellenaran la casilla si no la recordaban. La mayor eficiencia para los diagramas con estereotipos parece confirmar esta interpretación: los sujetos responden sólo lo que recuerdan en un pequeño periodo de tiempo, pero sin considerar el resto y haciéndolo de manera errónea o incompleta.
- **Transferencia:** los resultados fueron los contrarios a los de la Retención. Se obtuvieron mejores resultados para la efectividad cuando los sujetos usaron los diagramas con estereotipos, pero fueron peores para la eficiencia. Además, estos resultados confirman que

los diagramas con estereotipos resultan más comprensibles y menos propensos a error pero la cantidad de información que incluyen implica que es necesario más tiempo para el procesamiento del diagrama: en este caso, y de forma diferente que para la comprensión semántica, los sujetos tuvieron que reconstruir los flujos de los mensajes en el diagrama de secuencia y por ello tuvieron que emplear más tiempo para entender primero la semántica y después la sintaxis de los diagramas. Esto conllevó a emplear más tiempo y por ello menor eficiencia. Al igual que en la retención, los sujetos interpretaron mejor los diagramas del dominio B (ver los gráficos de cajas que se incluyen como ejemplo en la Figura 2).

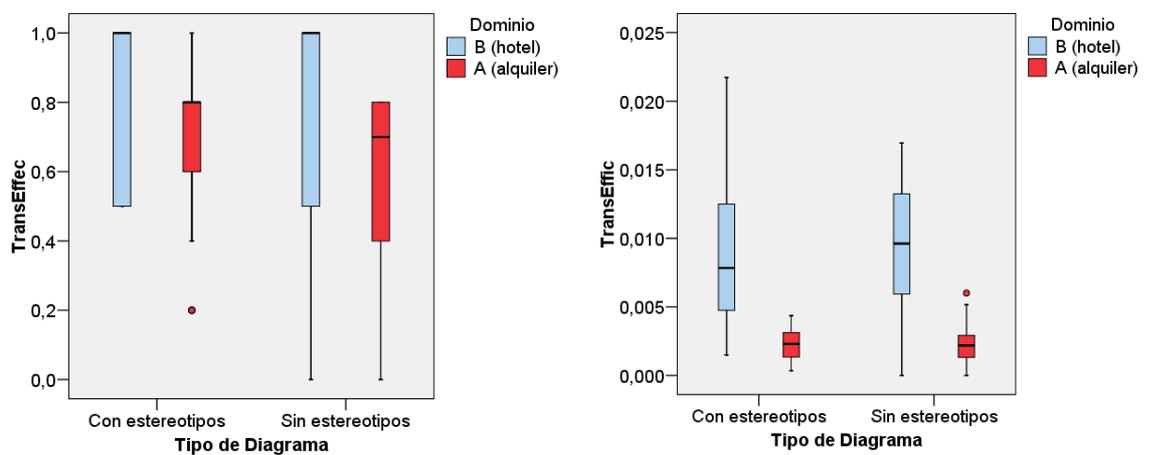


Figura 2. Diagrama de cajas para TransEfec (a) y TransEfic (b)

La Tabla 3 presenta los resultados obtenidos de los ANOVAs realizados para comprobar las hipótesis presentadas en la sección 4.1. En esta tabla, podemos observar que en mucho de los casos, el dominio afecta significativamente la entendibilidad de los diagramas (ver las celdas en negrita).

Tabla 3. Resultados de los ANOVAs

Efectividad		CSEfec			RetenEfec			TransEfec		
Origen	gl	F	sig.	Potencia	F	sig.	Potencia	F	sig.	Potencia
Estereotipos	1	0,749	0,388	0,138	0,206	0,650	0,074	0,798	0,373	0,144
Dominio	1	0,018	0,893	0,052	35,369	0,000	1,000	17,680	0,000	0,987
Eficiencia		CSEfic			RetenEfic			TransEfic		
Origen	gl	F	sig.	Potencia	F	sig.	Potencia	F	sig.	Potencia
Estereotipos	1	0,311	0,578	0,086	0,126	0,723	0,064	0,157	0,693	0,068
Dominio	1	9,422	0,003	0,862	3,386	0,068	0,448	98,538	0,000	1,000

La Tabla 2 mostraba que los sujetos lo habían hecho peor con los diagramas del dominio A (alquiler de coches). Esto puede indicar que el dominio del alquiler de coches fue más difícil para los sujetos. El gráfico de perfil de la Figura 3 muestra, a modo de ejemplo, este efecto para la medida de CSEfic.

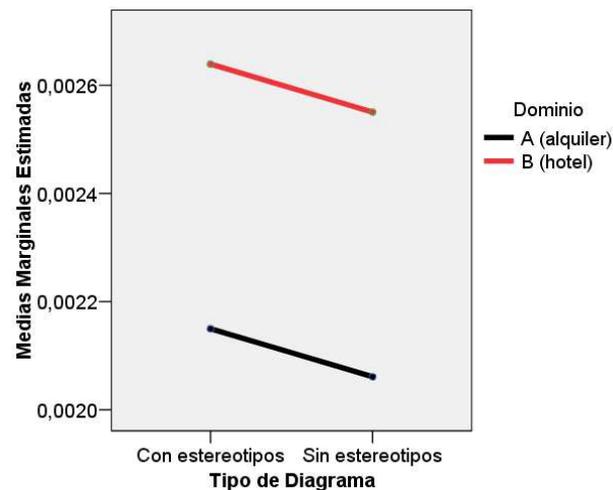


Figura 3. Gráfico de perfil de CSEfic

Con respecto al uso de estereotipos, no existen valores significativos en la Tabla 3. Además, la potencia de los tests es baja. Esto implica que no podemos rechazar la hipótesis nula investigada en este trabajo mientras que aceptar la hipótesis alternativa puede ser arriesgado y engañoso. Además, tenemos resultados que pudieron ser consecuencia del tratamiento o determinados por un efecto de interacción entre el tratamiento y el dominio.

5.1. Amenazas a la validez del experimento

Considerando algunas cuestiones que pueden amenazar la validez del experimento, quizás la más discutible pueden ser las amenazas a la validación externa, que conciernen a la generalización de las conclusiones.

Este tipo de amenazas se presentan siempre cuando experimentamos con alumnos y, en este caso, somos conscientes de que no tienen un conocimiento previo sobre los estereotipos usados en los diagramas de secuencia UML y que se incluyen en el material del experimento. Además, incluso si el conocimiento de los alumnos pudiera asumirse como comparable a un profesional junior de la industria, la presión del trabajo y el ambiente global son distintos. Otra amenaza para la validez de las conclusiones de este trabajo es la asunción de la formación homogénea entre los alumnos, que no se ha evaluado antes de la ejecución del experimento.

Finalmente, la hipótesis de equivalencia entre dominios no fue confirmada por el análisis de datos y esto ha llegado a ser uno de los obstáculos de este trabajo. De hecho, existieron diferencias relevantes en el rendimiento de los alumnos que pueden ser atribuidas a las diferencias entre los dominios de los diagramas: el diagrama del alquiler de coches resultó más difícil que el del hotel.

Para controlar estas amenazas estamos planificando realizar replicas de este experimento.

6. Conclusiones

Este trabajo ha analizado el uso de estereotipos en el contexto de los diagramas de secuencia. Se ha realizado un experimento controlado y los resultados obtenidos indican que, en general, los diagramas estereotipados pueden ser más comprensibles, menos propensos a errores y requerir menos tiempo para ser entendidos que los diagramas que no los utilizan. En cuatro de los seis casos estudiados, la puntuación obtenida por los alumnos fue superior cuando usaron estereotipos.

Sin embargo, estos resultados no pueden ser asumidos como definitivos. Los p-valores obtenidos en los tests de hipótesis no son estadísticamente significativos y, por lo tanto, las hipótesis nulas no pueden ser rechazadas. Esto implica que, aunque estos resultados indiquen una mejora de la entendibilidad con el uso de estereotipos, esto no se puede confirmar estadísticamente.

Estos resultados favorecen otras investigaciones que se pretenden realizar en un futuro próximo, tratando de controlar las amenazas de una forma más efectiva a través de cuestiones como: (i) eliminar el efecto dominio; (b) tener en cuenta el factor de la experiencia intentando señalar el efecto que tiene en el rendimiento de los alumnos usando los estereotipos. Una posible forma podría ser evaluar la experiencia antes del experimento y consecuentemente realizar los grupos basándose en dicha experiencia; (c) mejorar el análisis de datos clarificando la diferencia entre las preguntas no contestadas y las preguntas erróneas y su impacto en la efectividad y en la eficiencia. Además, otros autores [3] sugieren investigar la posible existencia de otros factores relevantes, como la habilidad y la experiencia de los sujetos. Como trabajo futuro, se pretende replicar este experimento en otros contextos tanto académicos como industriales para determinar de forma concluyente si los estereotipos propuestos mejoran de forma significativa la entendibilidad de los diagramas de secuencia UML.

Agradecimientos

Este trabajo ha sido financiado por los siguientes proyectos: “META (TIN2006-15175-C05-05)” del Ministerio de Educación y Ciencia y los proyectos “IDONEO (PAC08-0160-6141)” y “EVVE (HITO-2008-49)” de la Consejería de Ciencia y Tecnología de la Junta de Comunidades de Castilla-La Mancha, el “ISECOMO (TC20091092)” de la Universidad de Castilla-La Mancha y el proyecto “Transformación de Modelos Dirigida por Atributos de Calidad” de la UPV.

Referencias

1. OMG: UML 2.1 Unified Modeling Language (TM) (2006)
2. Staron, M., Kuzniarz, L., Wohlin, C.: Empirical Assessment of Using Stereotypes to Improve Comprehension of UML Models: a Set of Experiments. *The Journal of Systems and Software* 79, 727-742 (2006)
3. Ricca, F., Di Penta, M., Torchiano, M., Tonella, P., Ceccato, M.: The Role of Experience and Ability in Comprehension Tasks supported by UML Stereotypes. In: Proc. of 29th In. Conference on Software Engineering (ICSE'07), pp. 375-384 (2007)
4. Genero, M., Manso, M.E., Visaggio, A., Canfora, G., Piattini, M.: Building Measure-Based Prediction Models for UML Class Diagram Maintainability. *Empirical Software Engineering* 12, 517-549 (2007)

5. Otero, M.C. Dolado, J.J.: Evaluation of the Comprehension of the Dynamic Modeling in UML. *Information and Software Technology* 46 (1),35-53 (2004)
6. Reinhartz-Berger, H. Dori, D.: OPM vs. UML - Experimenting with Comprehension and Construction of Web Application Models. *Empirical Software Engineering*, 10, 57-79 (2005)
7. Selic, B.: The Pragmatics of Model-Driven Development. *IEEE Software* 20(5). 19-25 (2003)
8. Insfran, E.: A Requirements Engineering Approach for Object-Oriented Conceptual Modeling, in DSIC. University of Technology of Valencia (2003)
9. Insfran, E., Pastor, P., Wieringa, R.: Requirements Engineering-based Conceptual Modeling. *Journal of Requirements Engineering* 7(2), 61-72, (2002)
10. Genero, M. Cruz-Lemus, J. A. Caivano, D. Abrahão, S. M. Insfrán, S., Carsí, J. A.: Does the use of stereotypes improve the comprehension of UML sequence diagrams? In: Proc. of the International Symposium on Empirical Software Engineering and Measurement (ESEM 2008) pp. 300-302 (short paper) (2008)
11. Cruz-Lemus. J.A., Genero, M., Manso, M^ºE., Morasca, S., Piattini, M.: Assessing the Understandability of UML Statechart Diagrams with Composite States - A Family of Empirical Studies. *Empirical Software Engineering* (por aparecer - DOI: 10.1007/s10664-009-9106-z) (2009)
12. Glezer, C., Last, M., Nachmany, E., Shoval, P.: Quality and Comprehension of UML Interaction Diagrams - An Experimental Comparison. *Information and Software Technology* 47, 675-692 (2005)
13. Lange, C.F.J., Chaudron, M.R.V.: Interactive Views to Improve the Comprehension of UML Models - An Experimental Validation. In: Proc. of 15th IEE30E Int. Conference on Program Comprehension (ICPC'07), pp. 221-2130. Canada (2007)
14. Purchase, H.C., Colpoys, L., McGill, M., Carrington, D.: UML Collaboration Diagram Syntax: an Empirical Study of Comprehension. In: Proc. of 1st Int. Workshop on Visualizing Software for Understanding and Analysis (VISSOFT'02), pp.13-22. Paris, France (2002)
15. Purchase, H.C., Colpoys, L., McGill, M., Carrington, D., Britton, C.: UML Class Diagram Syntax: an Empirical Study of Comprehension. In: Proc. of Australian Symposium on Information Visualisation, pp. 113-120. Sydney, Australia (2001)
16. Xie, S., Kraemer, E., Stirewalt, R.E.K.: Empirical Evaluation of a UML Sequence Diagram with Adornments to Support Understanding of Thread Interactions. In: Proc. of 15th IEEE Int. Conference on Program Comprehension (ICPC'07), pp. 123-134. Banff, Canada (2007)
17. Yusuf, S., Kagdi, H., Maletic, J.I.: Assessing the Comprehension of UML Class Diagrams via Eye Tracking. In: Proceedings of 15th IEEE International Conference on Program Comprehension (ICPC'07), pp. 113-122. Banff, Canada (2007)
18. Jacobson, I., Christerson, M., Jonsson, P., Overgaard, G.: *Object-Oriented Software Engineering, a Use Case-driven Approach*, Addison-Wesley, Reading, USA (1992)
19. Eriksson, H.E., Penker, M.: *UML Toolkit*, John Wiley and Sons (1998)
20. Larman, C.: *Applying UML and Patterns*, Prentice-Hall (1998)
21. Rolland, C., Ben Achour, C., Cauvet, C., Ralyté, J., Sutcliffe, A., Maiden, N.A.M., Jarke, M., Haumer, P., Pohl, K., Dubois, E., Heymans, P.: A Proposal for a Scenario Classification Framework. *Requirements Engineering* 3(1), 23-47 (1998)

22. Basili, V., Rombach, H.: The TAME Project: Towards Improvement-oriented Software Environments. *IEEE Trans. on Software Eng.* 14(6), 758-773 (1988)
23. Basili, V., Weiss, D.: A Methodology for Collecting Valid Software Engineering Data. *IEEE Transactions on Software Engineering* 10(6), 728-738 (1984)
24. Juristo, N., Moreno, A.: *Basics of Software Engineering Experimentation*, Kluwer Academic Publishers (2001)
25. Wohlin, C., Runeson, P., Hast, M., Ohlsson, M.C., Regnell, B., Wesslen, A.: *Experimentation in Software Engineering: an Introduction.*, Kluwer Academic Publisher (2000)
26. Kirk, R.E.: *Experimental Design: Procedures for the Behavioral Sciences*, Brooks/Cole Publishing Company (1995)
27. Bodart, F., Patel, A., Sim, M., Weber, R.: Should Optimal Properties Be Used in Conceptual Modelling? A Theory and Three Empirical Tests. *Information Systems Research* 12(4), 384-405 (2001)
28. Gemino, A., Wand, Y.: Complexity and Clarity in Conceptual Modeling: Comparison of Mandatory and Optional Properties. *Data and Knowledge Engineering*, 55, 301-326 (2005)
29. Mayer, R.E.: Models for Understanding. *Review of Educational Research* 59(1), 43-64 (1989)
30. Mayer, R.E.: *Multimedia Learning*, Cambridge University Press (2001)
31. SPSS: *SPSS 12.0, Syntax Reference Guide.*, SPSS Inc.: Chicago, USA (2003)
32. Winer, B.J., Brown, D.R., Michels, K.M.: *Statistical Principles in Experimental Design*, McGraw-Hill (1991)