

Influencia de las nominalizaciones sobre la completitud de modelos de requisitos

Claudia S. Litvak^{1,2}, Graciela D. S. Hadad¹, Jorge H. Doorn^{1,2}

¹Escuela de Informática, UNO

Córdoba 1055. Merlo. Buenos Aires. (0220) 482-0799

²Departamento de Ingeniería e Investigaciones Tecnológicas, UNLaM

e-mail: claudialitvak@gmail.com, gracielahadad@gmail.com, jdoorn@exa.unicen.edu.ar

Resumen

Dentro de la Ingeniería de Software surge como área altamente relevante, el estudio de los requisitos de un sistema de software. Definir todos los requisitos de un sistema en un contexto determinado es prácticamente imposible. Meramente estimar el grado de completitud logrado es un problema muy complejo. Para abordar el problema de la completitud se suelen generar modelos, que son validados con clientes y usuarios, siendo de particular interés los modelos en lenguaje natural, ya que facilitan la comprensión por parte del usuario. Pero aún contando con esta facilidad de comprensión por parte del usuario y contando con ingenieros de requisitos expertos, el problema de completitud de los requisitos desarrollados es notorio y debe ser atendido. En el presente trabajo se propone considerar específicamente las ventajas e inconvenientes que acarrea el uso de las nominalizaciones de los verbos en los modelos construidos, ya que se supone que un uso poco cuidado de las mismas puede desencadenar omisiones significativas.

Palabras clave: Ingeniería de Requisitos, Completitud de Modelos, Nominalizaciones.

Contexto

Esta propuesta está incluida en los proyectos “Tratamiento de los factores situacionales y la completitud en la ingeniería de requisitos” de la Universidad Nacional del Oeste (UNO) y “Control de Calidad de requisitos escritos en Lenguaje Natural” de la Universidad Nacional de la Matanza (UNLaM).

Introducción

El proceso de la Ingeniería de Requisitos tiene como primera actividad conocer cabalmente el dominio de la aplicación, a fin de definir correctamente los servicios que deberá prestar el sistema de software que está siendo desarrollado.

Usualmente esta actividad involucra el desarrollo de modelos que reflejan dicho dominio.

La utilización del lenguaje natural en el modelado del contexto y de los requisitos de un sistema software facilita la comunicación con todos los involucrados dado que es comprendido por todos, y de ese modo los clientes y

usuarios pueden participar más eficazmente en ese proceso.

El presente proyecto se encuentra enmarcado en un proceso de requisitos desarrollado a lo largo de años y que ha sido utilizado en numerosos casos de estudio, tanto a nivel académico de grado y postgrado como en el ámbito profesional [Leite97] [Leite00] [Leite04]. Los modelos presentados por Leite et al. [Leite97] son el Léxico Extendido del Lenguaje (LEL) y los Escenarios, modelos que han probado ser idóneos para facilitar la comunicación entre los involucrados. El modelo LEL es un glosario con términos que corresponden al contexto donde el software se implantará, mientras que los escenarios son descripciones de situaciones en el universo de discurso, es decir, en dicho contexto.

El vocabulario intrínseco a un dominio particular transporta información. Esto se debe a que la cultura del dominio de la aplicación se materializa en un alto grado en su vocabulario [Nettle00] [Fishman99]. Es por ello fundamental, como parte de las actividades destinadas a conocer el dominio de la aplicación, adaptarse lo más posible a ese vocabulario. Leite enfatiza [Leite01] que el proceso de definición de los requisitos es un proceso inherentemente incompleto, dada la complejidad del mundo real. Por ello plantea que al no poderse alcanzar la completitud es deseable obtener modelos “lo más completos posible”. Esto debe ocurrir con todos los modelos del proceso, comenzando por el modelo que inicialmente se construye: LEL.

Si bien estos conceptos han estado presentes desde las primeras etapas de definición del presente proceso de requisitos, en etapas avanzadas del desarrollo del mismo se ha hecho cada vez más evidente la importancia de la completitud de los modelos utilizados. El

proceso de construcción del modelo LEL ha sido ampliamente desarrollado en [Leite95] [Leite97] [Leite04], puntualizando en el estudio de la completitud de dicho modelo [Doorn03] [Ridao06] [Doorn08] [Hadad12] [Litvak12] [Hadad13] [Litvak13a] [Litvak13b] [Hadad14] [Litvak14] de modo de hallar la mayor cantidad de términos posibles de dicho glosario.

En la construcción del LEL, se define para cada término del LEL un nombre, con todos los sinónimos involucrados, su denotación y connotación. Luego se catalogan los términos símbolos según el tipo: sujeto, objeto, verbo o estado.

Con la experiencia de décadas trabajando con el LEL, utilizado en cursos de grado, cursos de posgrado y en aplicaciones reales, se han observado inconvenientes en el tratamiento de los verbos. Estos inconvenientes están relacionados principalmente con el uso de un término como verbo en su forma verbal o en su forma nominalizada.

La nominalización consiste en la construcción de sustantivos, a partir de verbos o adjetivos [Alexiadou01] [Azpiazu02] [Bisetto05] [NGRAE09]. El presente trabajo se aboca específicamente a las nominalizaciones de los verbos. La forma más sencilla de describir las nominalizaciones es mediante la frase *acción y efecto de*. Aunque en algunos casos las nominalizaciones son sólo la *acción de* o sólo el *efecto de*. El punto central es que no es lo mismo registrar en un glosario el verbo que sus nominalizaciones. Este aspecto es más notable a la hora de registrar los impactos. Un impacto describe la connotación del término del LEL, es decir, describe cómo el símbolo actúa en el dominio de la aplicación.

La primera idea sería confirmar que las nominalizaciones son en algunos casos más precisas que el verbo mismo,

en infinitivo o conjugado. Además, muchos verbos pueden producir más de una nominalización. Entonces queda por estudiar si realmente es específico del dominio de aplicación el verbo, el verbo más su nominalización (sustantivo obtenido o estado generado) o solamente la nominalización.

Líneas de Investigación, Desarrollo e Innovación

Se ha observado que al estudiar el mismo problema, diferentes personas utilizan el verbo mismo mientras que otras utilizan su nominalización. Esto es potencialmente una causa de ambigüedad, ya que el verbo en sí mismo puede arrastrar impactos del dominio de la aplicación diferentes de los que arrastra su versión nominalizada, especialmente cuando esta adquiere un significado de objeto o estado.

Posiblemente esto lleve a descubrir homónimos ocultos o a registrar por separado el verbo por un lado y el estado o el objeto producido por el otro. Seguramente, de ser este el caso se obtendrían una mayor cantidad de impactos y una mejora en la especificidad de los mismos.

A partir de todas estas razones ha surgido la presente línea de investigación, que aspira a mejorar la completitud del LEL mediante un mejor tratamiento de los verbos y sus nominalizaciones.

Resultados y Objetivos

Hasta la fecha se ha realizado un estudio del concepto de nominalización [Alexiadou01] [Azpiazu02] [Bisetto05] [NGRAE09] y se han avizorado las consecuencias que su uso no controlado es una posible fuente de deficiencias en el proceso de requisitos.

Las nominalizaciones se clasifican en dos grandes grupos: a) Eventivas y b) No eventivas. Las nominalizaciones Eventivas denotan la existencia de un proceso que ocurre o tiene lugar, de ahí que en la oración se combinen con predicados como *tener lugar*, *durar*, *ocurrir*, etc., o bien con adjuntos temporales sin la preposición *de*.

Las nominalizaciones No eventivas no indican un proceso. Pueden categorizarse en varios tipos. Uno de estos son las nominalizaciones de resultado, definidas en distintos trabajos a partir de la negación de las propiedades que caracterizan una nominalización Eventiva [Grimshaw90]. En segundo lugar, junto con las nominalizaciones de resultado, existen otras nominalizaciones No eventivas que no indican un resultado de la acción, sino que expresan diversos contenidos significativos, como objetos y estados, no resultantes de la acción significada en la base verbal, lugares o instrumentos.

La primera conclusión obtenida ha sido que las nominalizaciones son en algunos casos más precisas que el verbo en infinitivo y en otros igual de ambiguas. Son más precisas cuando su condición de Eventiva o No eventiva no es ambigua (ejemplo: agrupación/ agrupamiento), pero pueden ser también ambiguas cuando la condición de Eventiva o No eventiva depende del contexto (ejemplo: construcción). Eso es un aspecto importante en la reducción de la ambigüedad.

El objetivo del presente trabajo es estimar cómo influye en la completitud la presencia de la nominalización de un verbo, del verbo mismo o la presencia de ambas. Además, se pretende determinar si los impactos definidos pertenecen al verbo mismo o a su nominalización. Es posible que la migración de impactos entre verbos y sus nominalizaciones sea

una de las causas de la aparición de omisiones en el modelo LEL.

Asimismo, se aspira a continuar con el estudio de la completitud del LEL, avanzando hacia la completitud de los modelos posteriores como los Escenarios y el documento de Especificación de Requisitos de Software.

Esto último incluye mejorar los resultados ya logrados acerca de la influencia de la completitud del LEL sobre los modelos posteriores.

Formación de Recursos Humanos

El LITIC de la Universidad Nacional del Oeste, fue creado en 2014, con investigadores de amplia trayectoria en otras universidades. A la fecha, el Proyecto de Ingeniería de Requisitos del LITIC cuenta con tres investigadores. La propuesta es agregar mayor cantidad de investigadores, formando a docentes de la Universidad Nacional del Oeste como investigadores. También se ha solicitado la incorporación de becarios para que participen en este proyecto.

En la Universidad Nacional de La Matanza, desde hace 10 años existe un grupo de investigación del área de Ingeniería de Requisitos, perteneciente al Departamento de Ingeniería e Investigaciones Tecnológicas. Una de las ramas actuales de investigación de este grupo es la completitud. En particular actualmente hay integrantes del grupo avocados al estudio de las nominalizaciones, sobre dos casos prácticos específicos.

Además, este trabajo forma parte de la Tesis de Doctorado que la Magíster Claudia Silvia Litvak está desarrollando en UNLP, la cual trata sobre la completitud de los modelos de requisitos y se denomina “Gestión de la completitud en la Ingeniería de Requisitos”.

Referencias

[Alexiadou01] Alexiadou, A. (2001). *Functional Structure in Nominals. Nominalization and ergativity*. Amsterdam, The Netherlands. John Benjamins Publishers, ISBN 9789027227638.

[Azpiazu02] Azpiazu S. (2002). *Las estrategias de nominalización y el adverbio: estudio contrastivo del caso español*. Salamanca. Ediciones Universidad de Salamanca, ISBN 8478007873.

[Bisetto05] Bisetto A., Melloni C. (2005). *Result Nominals: a Lexical-Semantic Investigation*. Geert Booij et al. (eds.), *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5)*, Fréjus 15-18 September 2005, University of Bologna, pp. 393-412.

[Doorn03] Doorn, J.H., Ridao, M. (2003). *Completitud de Glosarios: Un Estudio Experimental*. 6th Workshop on Requirements Engineering Paracicaba, Brasil: Universidade Metodista de Piracicaba, pp. 317-328.

[Doorn08] Doorn, J.H., Ridao, M.N. (2008). *Completeness Concerns in Requirement Engineering*. In Mehdi Khosrow-Pour (Ed.) *Encyclopedia of Information Science and Technology*, Second Edition Hershey, PA: IGI Global, pp. 619-624.

[Fishman99] Fishman, J. (1999). *Handbook of Language and Ethnic Identity*. Oxford: Oxford University press. ISBN10: 0195374924.

[Grimshaw90] Grimshaw, J. (1990). *Argument structure*. *Linguistic Inquiry Monographs* 18. Cambridge, MIT Press. ISBN10: 0262570904.

[Hadad12] Hadad G.D.S., Litvak C.S., Doorn J.H. (2012). *Agregando semántica a técnicas de predicción de completitud en modelos de requisitos*. WICC 2012 - XIV Workshop de Investigadores en Ciencias de la Computación, anales electrónicos, Universidad Nacional de Misiones, Posadas, ISBN: 978-950-766-082-5, pp. 392-396, <http://hdl.handle.net/10915/18863>.

[Hadad13] Hadad, G.D.S., Litvak, C.S., Doorn, J.H., Ridao M.N. (2013). *Dealing with Completeness in Requirements Engineering*. *Encyclopedia of Information Science and Technology*, Third Edition. Editorial: IGI Global, Mehdi Khosrow-Pour (ed), *Information ScienceReference*, Hershey, PA, EEUU, 3° edición, cap. 279, pp. 2854-2863, Julio 2014 (10384 p.). ISBN13: 9781466658882.

[Hadad14] Hadad G.D.S., Litvak C.S., Doorn J.H. (2014). Problemas y Soluciones en la Completitud de Modelos en Lenguaje Natural. CADI 2014 – II Congreso Argentino de Ingeniería, Universidad Nacional de Tucumán, San Miguel de Tucumán, ISBN: 978-987-1662-51-7, T366, Septiembre 2014. Artículo destacado mediante la aceptación para su exposición oral, además de su publicación en los anales electrónicos. <https://cadi.org.ar/index.php/trabajos-seleccionados/16-tecnologia-de-la-informacion-y-comunicacion>.

[Leite95] Leite, J.C.S.P., Oliveira, A.P.A. (1995). A Client Oriented Requirements Baseline. Second IEEE International Symposium on Requirements Engineering, IEEE Computer Society Press, pp. 108-115.

[Leite97] Leite, J.C.S.P., Rossi, G., Balaguer, F., Maiorana, V., Kaplan, G., Hadad, G., Oliveros, A. (1997). Enhancing a Requirements Baseline with Scenarios. Requirements Engineering Journal, Springer-Verlag London Ltd., Vol. 2, N° 4, pp. 184-198.

[Leite00] Leite J.C.S.P., Hadad G.D.S., Doorn J.H., Kaplan G.N. (2000). A Scenario Construction Process. Requirements Engineering Journal, ISSN: 0947-3602, Vol.5, N° 1, pp. 38-61, Springer-Verlag London Ltd., Londres, Reino Unido, Julio 2000. <http://link.springer.com/article/10.1007/PL00010342>

[Leite01] Leite, J.C.S.P. (2001). Gerenciando a Qualidade de Software com Base em Requisitos. Qualidade de Software: Teoria e Prática. Prentice-Hall, Rocha A, Maldonado J, Weber K (eds), Cap. 17, pp. 238-246.

[Leite04] Leite, J.C.S.P., Doorn, J.H., Kaplan, G.N., Hadad, G.D.S., Ridao, M.N. (2004). Defining System Context using Scenarios, en el libro “Perspectives on Software Requirements”, Kluwer Academic Publishers, EEUU, ISBN: 1-4020-7625-8, Cap. 8, pp. 169-199.

[Litvak12] Litvak C.S., Hadad G.D.S., Doorn J.H. (2012). Un abordaje al problema de completitud en requisitos de software. CACIC 2012 – XVIII Congreso Argentino de Ciencias de la Computación, Universidad Nacional del Sur, Bahía Blanca, ISBN: 978-987-1648-34-4, pp. 827-836, Octubre 2012. <http://hdl.handle.net/10915/23715>.

[Litvak13a] Litvak, C.S., Hadad, G.D.S., Doorn, J.H. (2013). Correcciones semánticas en métodos de estimación de completitud de modelos en lenguaje natural. 16th Workshop on Requirements Engineering Montevideo, Uruguay: ORT University, pp. 105–117.

[Litvak13b] Litvak, C.S., Hadad, G.D.S., Doorn, J.H. (2013). Mejoras semánticas para estimar la completitud de modelos en lenguaje natural. CoNaIISI 2013, 1er Congreso Nacional en Ingeniería Informática / Sistemas de Información. Córdoba, Argentina: Universidad Tecnológica Nacional.

[Litvak14] Litvak C.S., Hadad G.D.S., Doorn J.H. (2014). Heurísticas para el modelado de requisitos escritos en lenguaje natural. CACIC 2014 - XX Congreso Argentino de Ciencias de la Computación, Universidad Nacional de La Matanza, Buenos Aires, ISBN: 978-987-3806-05-6, Octubre 2014, <http://sedici.unlp.edu.ar/handle/10915/42337>.

[Nettle00] Nettle, D., Romaine, S. (2000). Vanishing Voices. Oxford: Oxford University Press. ISBN10: 0195152468.

[NGRAE09] REAL ACADEMIA ESPAÑOLA y ASOCIACIÓN DE ACADEMIAS DE LA LENGUA ESPAÑOLA (2009). Nueva gramática de la lengua española, Madrid, Espasa Libros.

[Ridao06] Ridao, M.N., Doorn, J.H. (2006). Estimación de Completitud en Modelos de Requisitos Basados en Lenguaje Natural. 9th Workshop on Requirements Engineering. Rio de Janeiro, Brasil: Pontificia Universidade Católica do Rio de Janeiro. pp. 151-158.