View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Servicio de Difusión de la Creación Intelectual

Semántica Oculta en Modelos de Requisitos

Marcela Ridao [(1)] Jorge H. Doorn [(1) (2)]

(1) INTIA, Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina
(2) Universidad Nacional de La Matanza, Argentina
e-mail: {mridao, jdoorn}@exa.unicen.edu.ar

Resumen

La Ingeniería de Requisitos (IR) es un área dentro de la Ingeniería de Software que tiene por objetivo mejorar la forma en que se comprenden y definen sistemas de software complejos. Los modelos construidos en el proceso de la Ingeniería de Requisitos han sido concebidos procurando maximizar su expresividad en relación con su propósito. Este proyecto se enmarca en una estrategia de la IR basada en dos modelos principales: el LEL (Léxico Extendido del Lenguaje) y los Escenarios. Estos modelos constituyen la base de una estrategia para obtener una especificación de requisitos de software. Sin embargo, estos modelos parecen contener más información que la que puede obtenerse en una primera lectura. En este artículo se propone, justamente, realizar nuevas lecturas de estos modelos para extraer esa información adicional. En particular, se propone obtener información sobre la estructura semántica de glosarios del proceso del negocio, a partir de su visualización por medio de grafos.

Palabras clave: Ingeniería de Requisitos, Glosarios, Trazado de Grafos, Métodos Dirigidos por Fuerzas.

Contexto

Esta línea de I/D se enmarca en el proyecto acreditado Bases de Datos y Procesamiento de Señales, que se desarrolla en la Facultad de Ciencias Exactas de la Universidad Nacional del Centro de la Provincia de Buenos Aires, INTIA (INvestigación en Tecnología Informática Avanzada).

Introducción

En los últimos años, se ha prestado especial atención a problemas tales como las redes organizacionales o redes sociales organizacionales, redes de referencias bibliográficas, redes de grupos de interés, etc. En este contexto, han adquirido importancia las disciplinas dedicadas al estudio de fenómenos donde el aspecto dominante es la complejidad estructural y no la complejidad esencial de los elementos involucrados en la estructura [1] [2].

Algunos de los modelos de la Ingeniería de Requisitos pueden ser estudiados desde el punto de vista estructural. En particular, el LEL puede ser visualizado bajo una óptica estructural, si se lo representa mediante un grafo, donde los símbolos sean representados mediante nodos, y las menciones a otros símbolos como arcos conectando dichos nodos. De esta manera, el LEL puede visualizarse como una suerte de red lingüística con una estructura claramente compleja. Esta estructura contiene información implícita empotrada, adicional a la información explícita que cada nodo almacena.

Construir y analizar el grafo de los símbolos del LEL se constituye así en una suerte de minería de información y de conocimiento, ya que se obtiene información sintáctica acerca de la estructura del LEL y al mismo tiempo se adquiere parte del conocimiento subyacente bajo esa estructura.

El LEL, como se verá más detalladamente en la siguiente sección, es una representación de los símbolos del lenguaje del dominio del problema, cuyo objetivo principal es que el ingeniero de requisitos conozca el lenguaje que habla el usuario, posponiendo la comprensión del problema a etapas posteriores de la estrategia de IR. Sin embargo, se puede suponer que si el Universo de Discurso contiene áreas de interés diferenciadas, los términos correspondientes a cada una de estas áreas deberían tener un mayor grado de acoplamiento que aquellos que corresponden a áreas diferentes.

Por lo anterior, en este trabajo se propone analizar el grafo construido a partir del LEL, con el fin de estudiar los mencionados agrupamientos.

LEL: Léxico Extendido del Lenguaje

La construcción de un vocabulario que capture la jerga usada por los expertos del dominio ha sido propuesta por distintos autores [3] [4]. De hecho, varias experiencias han mostrado que un glosario del vocabulario de los clientes-usuarios es, en si mismo, una fuente de información para elicitar información del Universo de Discurso [5] [6] [7] [8] [9].

En este trabajo, se estudiará un modelo de glosario en particular: el Léxico Extendido del Lenguaje (LEL). El LEL es una representación de los símbolos del lenguaje del dominio del problema, cuyo objetivo principal es que el ingeniero de requisitos conozca el lenguaje que habla el usuario, sin preocuparse por entender el problema [10] [11].

Este léxico se construye utilizando lenguaje natural y está compuesto por símbolos que pueden ser Sujetos (realizan acciones), Objetos (las acciones se realizan sobre ellos), Verbos (acciones del sistema) y Estados significativos del sistema [12].

Cada símbolo tiene uno o más nombres o frases que lo identifican y dos tipos de descripciones, la noción y el impacto. La noción describe la denotación de la palabra o frase. Indica quién, cuándo ocurre, qué procesos involucra, qué significado tiene el símbolo, etc. El impacto describe la connotación del símbolo, es decir, su repercusión en el sistema. Cada entrada puede contener una o más nociones y uno o más impactos.

Trazado de Grafos: Métodos dirigidos por fuerzas

La Teoría de Grafos tiene diversidad de aplicaciones. La representación mediante nodos y conexiones es usada para representar redes físicas como circuitos eléctricos, carreteras, moléculas orgánicas, y también interacciones menos tangibles

como relaciones sociológicas, bases de datos, o el flujo de control de un programa computacional [13] [14].

El Trazado de Grafos aplica topología y geometría para derivar representaciones de grafos en dos dimensiones. Básicamente, consiste en una representación gráfica del grafo en el plano, usualmente destinada a una visualización conveniente de ciertas propiedades del grafo en cuestión o de los objetos modelados [15] [16].

En este trabajo, al representar el grafo de un LEL, el énfasis se pone en la estructura de la red, y no en los criterios estéticos utilizados generalmente en el trazado de grafos [17], como distribución uniforme de los nodos, minimización de cruces de arcos, uniformidad en la longitud de los arcos, simetría, etc.

Existen diferentes estilos de representación, adecuados a diferentes tipos de grafos o diferentes propósitos de representación [18]. Entre una gran variedad de algoritmos se destaca una familia de métodos conocidos como "dirigidos por fuerzas". Estos métodos son muy usados hoy en día para dibujar grafos, porque dan buenos resultados, son sencillos de implementar, y son muy flexibles, por lo que pueden ser fácilmente adaptados a aplicaciones concretas con requerimientos de visualización específicos [19] [20]. Estos algoritmos usan analogías físicas para dibujar el grafo. Tienen como denominador común las siguientes características:

- Modelan al grafo como un sistema físico.
- El trazado del grafo es obtenido buscando el equilibrio del sistema físico.

Los modelos físicos más comunes son los que consisten de un sistema de fuerzas (donde generalmente se definen fuerzas que actúan entre los vértices del grafo), en cuyo caso el objetivo del algoritmo es encontrar una posición para cada vértice, de manera que el total de la fuerza ejercida en cada vértice sea cero.

Entre los primeros autores aplicando analogías con sistemas físicos para el trazado de grafos, se destaca el "Spring Embedder" propuesto por Eades [21], que se basa en reemplazar los nodos por anillos de acero y cada arco con un resorte para formar un sistema físico. Los nodos son ubicados en alguna disposición inicial, y se dejan actuar las fuerzas de los resortes hasta lograr un estado de energía mínima. La implementación de Eades, sin embargo, no siguió al pie de la letra la ley de Hooke, sino que incorporó al cálculo de las fuerzas resultantes, fuerzas repulsivas calculadas entre los nodos no conectados.

Otros autores proponen algoritmos derivados del Spring Embedder de Eades, como Kamada y Kawai [22], Davidson y Harel [23], Chernobelskiy et al [24], Kobourov [25], Bannister et al [26], Fruchterman y Reingold [17], entre otros.

En este trabajo, se utilizará el algoritmo propuesto en [17], que se basa en los siguientes principios:

- Los nodos conectados por un arco deberían ser dibujados cerca.
- Los nodos no deberían ser dibujados *demasiado* cerca uno de otro.

Cuán cerca se deberían ubicar los nodos, depende de cuántos haya y cuánto sea el espacio disponible. El algoritmo se basa en simulaciones moleculares o planetarias. Si los nodos se comportan como partículas atómicas o cuerpos celestes, ejerciendo fuerzas atractivas y repulsivas sobre los demás, las fuerzas inducen movimiento. Sin embargo, no se propone una simulación exactamente fiel a la realidad. Del mismo modo que en el algoritmo de Eades, sólo los nodos que son vecinos se atraen entre si, mientras todos los vértices se repelen unos a otros. Esto es consistente con la asimetría propuesta por los dos principios antes enunciados.

Objetivos

Con el fin de detectar agrupamientos de símbolos, se propone aplicar el algoritmo propuesto por Fruchterman y Reingold [17] en la visualización de los grafos correspondientes a los Léxicos de diferentes casos de estudio.

Para un LEL dado, entonces, se propone construir un grafo donde los símbolos son los nodos y las menciones a otros símbolos, los arcos.

Para la configuración inicial, los nodos se ubican al azar en el marco de trabajo, y luego se aplica el algoritmo, modificando la ubicación de los nodos en forma iterativa. Cada iteración consta de tres pasos:

- calcular el efecto de las fuerzas atractivas sobre cada nodo,
- calcular el efecto de las fuerzas repulsivas y
- modificar la posición de los nodos.

Inicialmente, se efectuaron pruebas utilizando el algoritmo con las fuerzas propuestas por Fruchterman y Reingold. En este trabajo, se propone modificar el sistema de fuerzas, comparando los resultados obtenidos.

Las fuerzas

f_a y f_r son las fuerzas de atracción y de repulsión, respectivamente.

Las pruebas iniciales, fueron efectuadas utilizando el sistema de fuerzas propuesto por [17]:

$$fa(d) = c1 * \frac{d^2}{k} \qquad k = C * \sqrt{\frac{Area}{NúmeroNodos}}$$

$$fr(d) = c2 * \frac{k^2}{d}$$

d es la distancia entre los vértices y k el radio vacío alrededor de un nodo

Posteriormente, se trabajó con el par de fuerzas propuesto por Eades [21]:

$$fa(d) = c3 * \log\left(\frac{d}{c4}\right)$$

$$fr(d) = c5 * \sqrt{d}$$

Las constantes C, c1, c2, c3, c4 y c5 son obtenidas experimentalmente.

Marco de trabajo

El grafo debe ser confinado al marco especificado por el usuario.

El algoritmo propuesto por [17] considera la ubicación de nodos ficticios en el perímetro del marco de trabajo, que ejercen fuerzas repulsivas sobre los nodos del grafo, pero ellos mismos permanecen fijos. De este modo, el marco se modela como cuatro paredes que contienen al grafo dentro de ellas.

Sin embargo, las experiencias realizadas permitieron determinar que no era necesaria la ubicación de nodos en las paredes del marco de trabajo, ya que, con la elección

adecuada de las constantes que intervienen en el cálculo de las fuerzas, los nodos se mantienen dentro del marco definido.

Resultados obtenidos

Con el fin de verificar si el algoritmo propuesto permite detectar agrupamientos de símbolos, se lo aplicó a diferentes casos de estudio cuyos Léxicos habían sido verificados y validados previamente.

Caso 1. ***Sistema de Planes de Ahorro Previo para la Adquisición de Vehículos 0Km*** [27]

Para este caso se prefirió no analizar en forma semántica la presencia o ausencia de agrupamientos de símbolos, con el fin de comprobar si éstos eran detectados por el algoritmo.

En primera instancia, se aplicó el algoritmo, utilizando el primer sistema de fuerzas. En la figura 1a, se presentan los resultados obtenidos.

Luego, se aplicó el algoritmo con el segundo par de fuerzas al mismo caso de estudio, obteniendo los resultados que aparecen en la figura 1b.

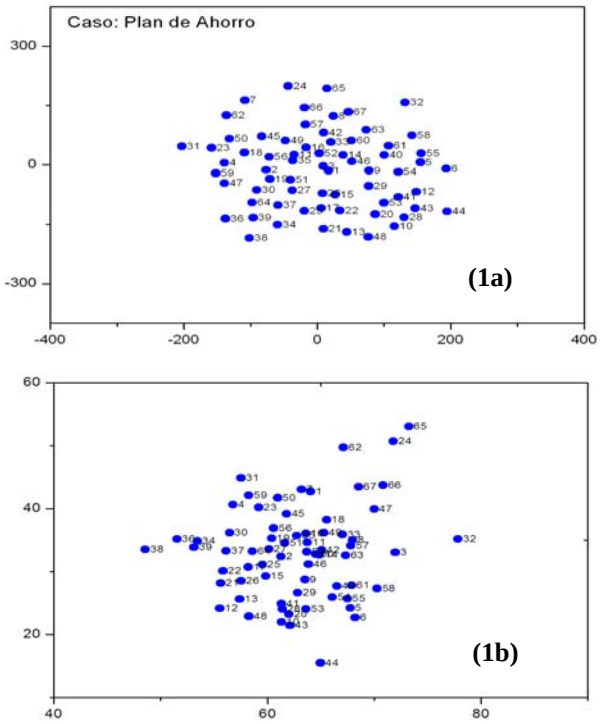

Fig. 1. Distribución de nodos caso 1

Cabe destacar que, tanto en éstas como en las siguientes figuras, no se muestran los arcos correspondientes a los vínculos entre los símbolos, para permitir una mejor visualización de la distribución de los nodos.

En el grafo presentado en la figura 1a no se observan agrupamientos. En la figura 1b, en cambio, se puede observar que algunos símbolos tienden a agruparse.

Un análisis semántico preliminar del léxico utilizado no mostró agrupamientos claramente definidos, aunque se distinguieron grupos de símbolos que podrían constituir clusters.

Para este caso de estudio, se contaba con los modelos de LEL y Escenarios desarrollados por 9 grupos de trabajo diferentes. Los grafos presentados en la figura 1 corresponden al LEL de uno de esos grupos.

Los modelos creados por cada grupo difieren en cuanto a los símbolos y Escenarios creados, en mayor o menor grado. Sin embargo, existen áreas de interés en el Universo de Discurso estudiado, que la mayoría de los grupos observó.

Por ello, se planea efectuar un análisis semántico detallado de los Léxicos disponibles para este caso de estudio, para comprobar si los agrupamientos que se observan en los grafos, efectivamente existen en el Universo de Discurso.

Caso 2. ***LEL del proceso de Construcción de LEL y Escenarios*** [28].

Para este caso de estudio, se conocía de antemano la existencia de clusters. En este caso, existen al menos dos agrupamientos: uno constituido por los símbolos correspondientes a la Construcción del LEL, y otro constituido por los símbolos correspondientes a la construcción de Escenarios.

En la figura 2a se presenta la distribución de los nodos obtenida utilizando el primer conjunto de fuerzas.

Se verifica que todos los símbolos que corresponden a la construcción del LEL se agrupan a la izquierda del grafo, mientras los que corresponden a la construcción de Escenarios se agrupan a la derecha. En la zona central, se ubican aquellos símbolos relacionados tanto con un proceso como con el otro. Por ejemplo, el nodo 8 corresponde al símbolo Cliente-Usuario, y el nodo 49 corresponde al símbolo Contexto del Problema.

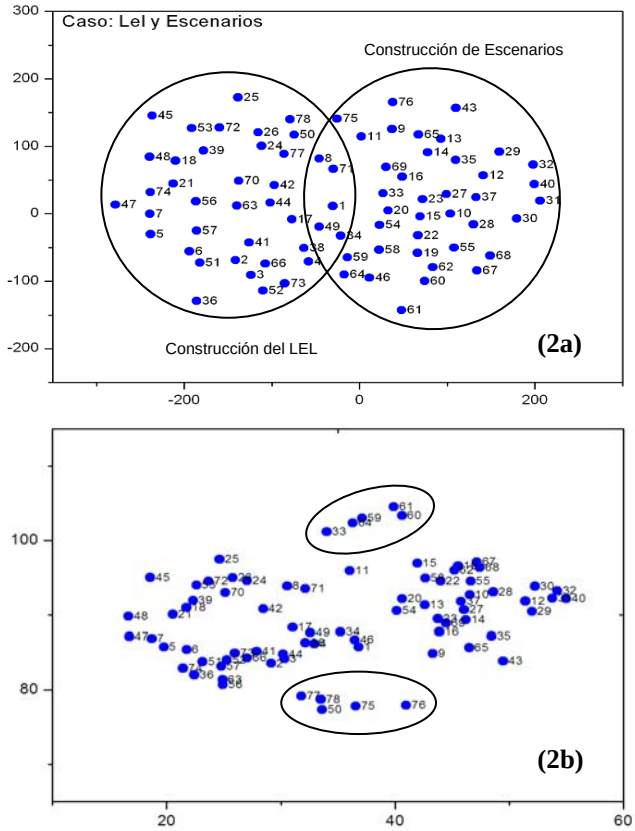

Fig. 2. Distribución de nodos para el caso 2

En el grafo de la figura 2b, obtenido con el segundo sistema de fuerzas, se observa también que los símbolos correspondientes a la construcción del LEL tienden a agruparse de un lado, mientras los correspondientes a la construcción de escenarios, lo hacen al otro. Pero, en este caso, se distinguen además otros agrupamientos más

pequeños, que no se observaban en la figura anterior. Cada uno de estos grupos fue comprobado en el análisis semántico del caso de estudio.

A modo de ejemplo, en la figura 2b se marcaron dos de los agrupamientos, y en la tabla 1 se presentan los nombres de los símbolos pertenecientes a los mismos.

Tabla 1. Símbolos correspondientes a dos de los clusters visibles en la figura 2b.

Vista de Hipertexto en los modelos de LEL y Escenarios	Extracción de los Requisitos del Sistema desde los Escenarios
50 Navegación	33 Escenario
75 Vínculo	59 Requisito del sistema
76 Vínculo a escenario	60 Requisito Funcional
77 Vínculo a símbolo del LEL	61 Requisito no Funcional
78 Vínculo arbitrario	64 Sistema de Software

Caso 3. *Relación entre productores de papa y una empresa acopiadora.*

Un análisis semántico preliminar de este caso de estudio, determinó la existencia de dos agrupamientos de símbolos; un grupo correspondiente a Producción y Entrega de Papas y otro a Canje de Semillas.

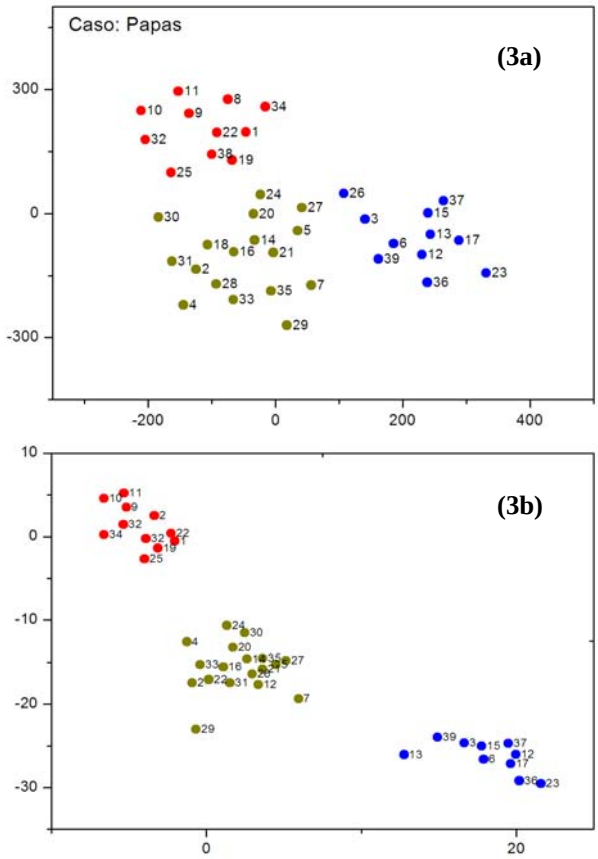

Fig. 3. Distribución de nodos para el caso 3

Sin embargo, en la visualización del grafo luego de la aplicación del algoritmo, tanto con el primer sistema de fuerzas, como con el segundo, se observa claramente la presencia de tres agrupamientos, como puede verse en las figuras 3a y 3b respectivamente.

Con el fin de detectar los límites de los clusters visualizados, se aplicó el clasificador *K*-Means [29] a los datos obtenidos, indicando la presencia de 3 grupos. En la figura 3 se muestran los clusters detectados con diferentes colores, y en la tabla 2, se presentan algunos de los símbolos pertenecientes a cada agrupamiento.

Tabla 2. Fragmento de la lista de símbolos correspondientes a cada uno de los clusters

Producción y Entrega
2 Bonificación
4 Causa justificada
5 Contrato de adquisición y producción
14 El Productor
16 Recibir papa
18 Establecer Programa de Entregas
20 Inspeccionar papa
21 MC S.A.
24 Pagar por la papa
…
Canje de Semilla
3 Cancelar la operación de canje
6 Contrato de canje de semilla
12 División de Semillas MC S.A.
13 El Comprador
15 El Vendedor
17 Entregar semillas
23 Orden de carga
26 Papa Grado MC Consumo
37 Semilla de Papa
…
Calidad
1 Agrietadura
8 Corazón hueco
9 Defecto de calidad
10 Defecto externo
11 Defecto interno
19 Grado MC Consumo
22 Malformación
25 Papa chica
38 Tolerancias establecidas
…

Un análisis semántico más profundo del caso de estudio, confirmó la presencia de estos tres agrupamientos. En particular, el tercer grupo, correspondiente a Calidad de la Papa, no había sido detectado en el análisis inicial, ya que los símbolos incluidos en él, se habían considerado parte del cluster correspondiente a Producción y Entrega de Papa.

Es importante destacar que estos resultados realzan la capacidad predictiva de la estrategia, ya que, más allá de confirmar la presencia de agrupamientos conocidos, permite descubrir la presencia de otros.

Conclusiones y trabajo futuro

Se ha propuesto un algoritmo que permite detectar agrupamientos de símbolos en un modelo del proceso de Requisitos: el Léxico Extendido del Lenguaje. El análisis de los resultados obtenidos hasta el momento, permite concluir que es posible efectuar una segunda lectura de algunos de los documentos de Ingeniería de Requisitos mediante el estudio de las estructuras semánticas subyacentes. El uso de grafos y el análisis de su estructura son herramientas apropiadas para dichos estudios.

Se ha comprobado que las fuerzas de atracción y repulsión que se apliquen a los nodos, inciden en la calidad de los resultados. Por ello, se propone considerar la aplicación de otros sistemas de fuerzas con el fin de verificar si es posible obtener una mejor visualización de los agrupamientos.

Los resultados obtenidos se acoplan en forma muy natural con estrategias genéricas de clusterización como K-Means. El

conocimiento previo del número de clusters es imprescindible para el uso de esta estrategia. Por ello, uno de los objetivos a lograr es la determinación del número de clusters detectados.

Entre otras posibilidades, se plantea estudiar la simetría existente en los grafos resultantes de la aplicación del algoritmo, así como la relación entre las variancias sobre los ejes X e Y. Esto permitiría descartar la presencia de clusters en aquellos casos donde exista simetría en los ejes X e Y, así como variancias similares sobre ambos ejes. Del mismo modo, la simetría o asimetría, según sea el caso, en forma conjunta con las variancias, podrían contribuir a la determinación del número de agrupamientos.

Se planea analizar, también, si la ubicación inicial de los nodos puede afectar la disposición final por el efecto que determinado nodo pudiera causar sobre sus vecinos, evitando su normal desplazamiento.

Además, se pretende aplicar el algoritmo a otras formas de representación del grafo. Por ejemplo, se planea estudiar la representación en tres dimensiones.

Se prevé avanzar en la correlación entre los resultados obtenidos mediante la aplicación de este algoritmo, con un análisis semántico detallado de cada uno de los casos de estudio. Para ello, se propone estudiar semánticamente los léxicos desarrollados por diferentes grupos de trabajo para un mismo caso de estudio, con el fin de detectar los agrupamientos semánticos detectados por cada grupo. Los resultados del algoritmo presentado serán un complemento muy útil para dicho estudio, ya que permitirán visualizar los agrupamientos que se observen en la lectura de los modelos.

Formación de Recursos Humanos

El grupo de investigación que participa en el proyecto Bases de Datos y Procesamiento de Señales está compuesto por 7 integrantes, todos ellos docentes de la Fac. de Ciencias Exactas categorizados en el Programa de Incentivos.

La línea de I/D forma parte de la tesis doctoral de la integrante del grupo Marcela Ridao, para el doctorado en Ciencias Informáticas de la Universidad Nacional de La Plata. El tema de la tesis es Técnicas Cuantitativas Orientadas al Reuso Semántico de Modelos de Requisitos.

Referencias

1. Barabasi, A.: Linked, The New Science of Network.. Perseus publishing (2002)
2. Dorogovtsev, S., Mendes, J.: Evolution of networks: From biological nets to the Internet and WWW. Oxford University Press, Oxford (2003)
3. Arango, G., Schafer, W., Prieto, R.: Domain Analysis Methods – Software Reusability. Ellis Horwood Ltd (1993)
4. Leite, J., Franco, A.: O Uso de Hipertexto na Elicitaçao de Linguagens da Aplicaçao. En: IV Simpósio Brasilero de Engenharia de Software, SBC, pp. 134-149. Brazil (1990)
5. Ben Achour, C., Rolland, C., Maiden, N., Souveyet, C.: Guiding Use Case Authoring: Results of an Empirical Study. In: International Symposium on Requirements Engineering, pp. 36-43. IEEE Computer Society Press, Limerick-Ireland (1999)
6. Rolland, C., Ben Achour, C.: Guiding the construction of textual use case specifications. Data & Knowledge Engineering 25, 125-160 (1998)
7. Oberg, R., Probasco, L, Ericsson, M.: Applying Requirements Management with Use Cases. Rational Software Corporation (1998)
8. Regnell, B.: Requirements Engineering with Use Cases – a Basis for Software Development, Doctoral Thesis. Department of Communication Systems. Lund University (1999)
9. Prakash, S., Aurum, A., Kox, K.: Requirements Engineering Practice in Pharmaceutical and Healthcare Manufacturing. In: 11th Asia-Pacific S.E. Conference, pp. 402-409 (2004)
10. Leite, J., Franco, A.: A Strategy for Conceptual Model Acquisition. In: IEEE International Symposium on RE, pp. 243-246. IEEE Computer Society Press (1993)
11. Leite, J., Rossi, G., Balaguer, F., Maiorana, V., Kaplan, G., Hadad, G., Oliveros, A.:. Enhancing a Requirements Baseline with Scenarios Requirements Engineeering Journal 2(4), 184-198 (1997)
12. Leite, J., Doorn, J., Kaplan, K., Hadad, G., Ridao, M.: Defining System Context Using Scenarios. In: Leite, J., Doorn, J (eds.) Perspectives on Software Requirements. Kluwer Academic Press, pp. 169-199 (2004)
13. Gross, J., Yellen, J.: Editors. Handbook of Graph Theory. CRC Press (2003)
14. Gross, J., Yellen, J.: Editors. Graph Theory and Its Applications, Second Edition (Discrete Mathematics and Its Applications). Chapman & Hall/CRC (2006)
15. Di Battista, G. , Eades, P. , Tamassia, R., Tollis, I.: Graph Drawing: Algorithms for the Visualization of Graphs. Prentice Hall (1999)
16. Brandes, U., Kenis, P., Wagner, D.: Communicating centrality in policy network drawing. IEEE Transactions on visualization and computer graphics 9(2), 241-253 (2003)
17. Fruchterman, T., Reingold, E.: Graph Drawing by Force-directed Placement. Software-Practice and Experience DOI 10.1002/spe.4380211102 Wiley Online Library (2006).
18. Kaufmann, M., Wagner, D.: (eds.) Drawing graphs: methods and models, LNCS, vol 2025. Springer-Verlag (2001)
19. Walshaw, C.: A multilevel algorithm for force-directed graph-drawing. Journal of Graph Algorithms and Applications 7(3), 253-285 (2003)
20. Aiello, A., Silveira, A.: Trazado de grafos mediante métodos dirigidos por fuerzas: revisión del estado del arte. Tesis de Licenciatura en Ciencias de la Computación (2004)
21. Eades, P.: A heuristic for graph drawing. Congressus Numerantium 42, 149–160 (1984)
22. Kamada, T., Kawai, S.: An algorithm for drawing general undirected graphs. Information Processing Letters 31, 7-15 (1989)
23. Davidson, R., Harel, D.: Drawing graphs nicely using simulated annealing, ACM Transactions on Graphics 15(4), 301-331 1996.
24. Chernobelskiy, R., Cunningham, K., Goodrich, M. T., Kobourov, S. G., Trott, L.: Force-directed Lombardi-style graph drawing, Proc. 19th Symposium on Graph Drawing, pp. 78–90. 2011
25. Kobourov, Stephen G.: Spring Embedders and Force-Directed Graph Drawing Algorithms, 2012.
26. Bannister, M. J., Eppstein, D., Goodrich, M. T., Trott, L.: Force-directed graph drawing using social gravity and scaling. Proc. 20th Int. Symp. Graph Drawing. 2012.
27. Rivero, L., Doorn, J., del Fresno, M., Mauco, V., Ridao, M., Leonardi, C.: Una Estrategia de Análisis Orientada a Objetos basada en Escenarios: Aplicación en un Caso Real. En: WER'98 - Workshop en Engenharia de Requisitos, pp. 79-90. Maringá, Brasil (1998)
28. Garcia, O., Gentile, C.: Diseño de una herramienta para contrucción de LEL y Escenarios, Graduation dissertation. Universidad Nacional del Centro de la Pcia. de Bs. As. (1999)
29. Peña, J., Lozano, J., Larrañaga P.: An empirical comparison of four initialization methods for the K-Means algorithm. Pattern Recognition Letters 20, 1027-1040 (1999)