

Modelagem de processos de negócio: efeito do método de notação no nível de ambiguidade

Business processes modeling: effect of notation method in the ambiguity level

Antônio Eduardo Carvalho e Silva¹
Rogerio Atem de Carvalho²
Hudson Silva Ferreira²

Como citar: Silva, A. E. C., Carvalho, R. A., & Ferreira, H. S. (2019). Modelagem de processos de negócio: efeito do método de notação no nível de ambiguidade. *Gestão & Produção*, 26(1), e1624. <https://doi.org/10.1590/0104-530X1624-19>

Resumo: A proposta deste estudo experimental controlado e aleatorizado foi analisar a especificação de requisitos funcionais. Avaliaram-se os possíveis impactos do método de notação de modelagem de processo no nível de ambiguidade presente na especificação em linguagem natural. As notações foram utilizadas como instrumento para manifestar as necessidades do usuário quanto ao processo de compra em aplicações de comércio eletrônico - caso de uso típico de sistemas de informação empresarial. A partir de um protótipo de baixa fidelidade, que representa o processo de compra numa loja virtual, 43 estudantes de graduação em ciência da computação foram agrupados de acordo com as notações desempenhadas: grupo controle linguagem natural (GC); grupo experimental máquina de estados finitos (GEMEF); e o grupo experimental notação combinada (GENC), que utilizou anotação manual de papéis semânticos baseados em máquina de estados finitos. Foi utilizado o conceito que trata da ambiguidade como informações inconsistentes que levam a múltiplas interpretações do documento de especificação dos requisitos funcionais. Pela técnica de leitura baseada em teste, associada à utilização da métrica de qualidade apropriada, foi conduzida uma análise de variância de fator único com delineamento completamente casualizado para saber se o método de notação, como fator primário, afeta o nível de ambiguidade. A comparação intergrupo sugere que: a notação combinada é o melhor método para reduzir o nível de ambiguidade da especificação de requisitos; e que a especificação de requisitos expressa em máquina de estados finitos gera o maior o nível de ambiguidade dentre as notações avaliadas. No entanto, estas tendências não são estatisticamente significativas. De forma geral, observou-se que a variável independente, método de notação, não afeta o nível de ambiguidade do processo descrito na especificação de requisitos funcionais.

Palavras-chave: Gestão de sistemas de produção; Processos de negócio; Especificação de requisitos; Nível de ambiguidade.

Abstract: The aim of this controlled and randomized experimental study was to analyze the functional requirement specification. The possible impacts of process modeling notation method in the ambiguity level of natural language specification were evaluated. The notations were used as a tool to express the user's needs about buying process in e-commerce applications - typical Enterprise Information System use case. From a low-fidelity prototype, that represents the purchasing process in a virtual store, 43 undergraduate computer science students were grouped according to the notations performed: natural (Brazilian Portuguese) language control group (CG); finite state machine experimental group (FSMEG); and combined notation experimental group (CNEG) that used manual annotation of semantic roles based on finite state machine. It was used the concept that deals with ambiguity as inconsistent information that leads to multiple interpretations of functional requirement specification document. By a Test-based Reading technique, associated with the use of appropriate quality metric, a one-way analysis of variance (completely randomized design) was conducted to know if notation method, as a primary factor, affects the ambiguity level. The inter-group comparison suggests that: combined notation is the best method to minimize ambiguity levels in requirement specification; and that requirement specification expressed in finite state machine generates the higher ambiguity level among assessed notations. However, these trends are not statistically significant. Generally, it was observed that the independent variable, notation method, does not affect the ambiguity level in the process described in the functional requirement specification.

Keywords: Management of production systems; Business processes; Requirements specification; Ambiguity level.

¹ Laboratório de Engenharia de Produção – LEPROD, Centro de Ciência e Tecnologia – CCT, Universidade Estadual do Norte Fluminense Darcy Ribeiro – UENF, Av. Alberto Lamego, 2000, Parque Califórnia, CEP 28013-602, Campos dos Goytacazes, RJ, Brasil, e-mail: eduardo.oak80@gmail.com

² Núcleo de Sistemas de Informação – NSI, Instituto Federal de Educação, Ciência e Tecnologia Fluminense – IFF, Rua Dr. Siqueira, 273, Parque Dom Bosco, CEP 28030-130, Campos dos Goytacazes, RJ, Brasil, e-mail: ratem@iff.edu.br; silvaferreira.hsf@gmail.com

Recebido em Set. 14, 2017 - Aceito em Maio 21, 2018

Suporte financeiro: Secretaria de Educação Profissional e Tecnológica do Ministério da Educação (SETEC/MEC).

1 Introdução

Como as necessidades do usuário podem ser traduzidas em código? Na metodologia ágil, as histórias utilizadas para especificar requisitos servem para nortear o desenvolvimento do sistema durante todo o seu ciclo de vida. Pela metodologia clássica, no entanto, a especificação de requisitos ocorre numa etapa inicial e bem definida do processo de desenvolvimento de sistemas (Leffingwell, 2011; Cohn, 2004). Independente da metodologia adotada para o desenvolvimento de sistemas, os requisitos são, via de regra, especificados numa linguagem carregada de ambiguidade – a linguagem natural (LN).

É possível analisar a ambiguidade da LN de duas perspectivas distintas. De um lado, a ambiguidade pode ser vista como um fator benéfico que motiva a dialética, assegurando o rigor do pensamento filosófico. No entanto, pela perspectiva do pensamento científico adotado no presente estudo, o fenômeno da ambiguidade foi analisado como um fator nocivo que afeta a especificação de requisitos (Piantadosi et al., 2012; Raffin, 2009; Walia & Carver, 2009; Sutcliffe et al., 1999; Card, 1998). Livrar-se da ambiguidade, neste contexto, representa um passo relevante para preencher a lacuna entre o modelo de negócio e a automação dos processos. O benefício da anotação semântica para esta finalidade é notória por tornar o processo passível de execução por máquina. Um exemplo disso é reportado no âmbito do *semantic business process modeling* (SBPM) no qual há uma integração entre modelo e metamodelo para formar uma ontologia que define os papéis semânticos (Yan, 2007; Lautenbacher et al., 2008).

O problema da ambiguidade intrínseca da LN pode dificultar o fluxo do conhecimento organizacional entre as partes envolvidas no processo. Essa lacuna da comunicação é especialmente nociva no desenvolvimento de um sistema de informação empresarial (SIE), pois, neste caso, há uma relação estreita entre o modelo lógico do negócio (MLN) e o código que provê o sistema. Portanto, à figura do usuário como parte interessada competem, de forma associada, os papéis de usuário do negócio e de usuário do sistema. Este papel complexo de usuário carece de meios que viabilizem sua integração ao processo de desenvolvimento (Coughlan & Macredie, 2002; Coughlan et al., 2003; Van Grondelle et al., 2010; Jørgensen et al., 2007; Carvalho et al., 2010, 2011, 2012).

A partir destas considerações que assumem a ambiguidade como um problema a ser superado, decorre o seguinte questionamento que motiva esta pesquisa: *como reduzir o nível de ambiguidade na especificação de requisitos funcionais prevista no processo de desenvolvimento de sistemas de informação empresariais?*

O objetivo geral desta pesquisa é analisar o efeito do método de notação de modelagem de processo de negócio no nível de ambiguidade da especificação de requisitos funcionais de SIEs. Os objetivos específicos foram: (i) identificar o perfil da distribuição dos dados

relativos ao nível de ambiguidade da especificação de requisitos funcionais de SIEs; (ii) propor um método combinado de notação para a modelagem de processos; (iii) aplicar o método combinado de notação na modelagem de processo de negócio; (iv) comparar o efeito dos métodos de notação de modelagem de processo de negócio no nível de ambiguidade da especificação de requisitos.

2 Métodos de notação para modelagem de processo de negócio

Sabe-se que a notação formal apresenta um potencial para a redução da ambiguidade inerente à LN e que a restrição imposta pelos autômatos finitos determinísticos (AFDs), como a máquina de estados finitos (MEF), reduz o potencial de descrição da LN (Lopes, 2008; Menezes, 2008a; Hopcroft et al., 2003). No entanto, é necessário criar mecanismos para extrair valor dos requisitos expressos em LN. Para que este valor seja devidamente reconhecido no universo dos sistemas de informação é necessário facilitar o fluxo de conhecimento entre os consumidores e outras partes interessadas no processo, estimulando o diálogo para aprimorar os processos organizacionais de forma contínua (Coughlan & Macredie, 2002; Coughlan et al., 2003; Van Grondelle et al., 2010; Jørgensen et al., 2007; Piantadosi et al., 2012; Leffingwell, 2011; Carvalho et al., 2010, 2011, 2012).

No âmbito dos SIEs, em que o papel do consumidor é empoderado pela figura do usuário do negócio e do usuário do sistema, sua participação ativa no processo produtivo é essencial na busca da missão do negócio (Kotler et al., 2012).

O desenvolvimento orientado ao comportamento (BDD), previsto na metodologia ágil, viabiliza a implementação direta de testes a partir das necessidades do usuário expressas pela tríade *Given-When-Then* (GWT). Neste íterim, a conexão direta “necessidade do cliente-sistema de informação (SI)” dispensaria a necessidade de formalizar o MLN com uma notação de modelagem. Na verdade, o MLN estaria implícito na tríade GWT prevista no BDD. Tal fato deve-se à possível analogia da tríade com a notação formal MEF, na qual *Given* seria um estado, *When* uma ação e *Then* o estado resultante. A tríade GWT caracterizaria, portanto, uma espécie de notação combinada (NC) que restringe a LN com a MEF disfarçada pela estrutura GWT. No entanto, a ambiguidade inerente à LN, utilizada para manifestar a necessidade do cliente, pode afetar a qualidade do fluxo de conhecimento seja na tradução que ocorre no eixo “MLN-SI” ou na conexão direta “necessidade do cliente-SI”. Assim, o desenvolvimento orientado à linguagem de negócios (BLDD), modelo conceitual que deriva do BDD, teve como proposta adicional a utilização da representação em grafo. Tal representação contribui para promover a conexão direta “necessidade do cliente-SI”, buscando

preservar a qualidade do fluxo de conhecimento e estreitar a relação descrição-código (Carvalho et al., 2010, 2011, 2012; Turnquist, 2011).

2.1 Linguagem natural

Devem-se assegurar meios que favoreçam a comunicação do usuário com os demais atores do processo de desenvolvimento do negócio/sistema. No entanto, o problema da ambiguidade da LN é um fator limitante para esta comunicação. O desafio em promover o envolvimento dos usuários na modelagem do conhecimento é representado pela falta de treinamento nas técnicas de representação formal do conhecimento por parte dos usuários (Van Grondelle et al., 2010; Carvalho et al., 2010, 2011, 2012).

ALN controlada é um método que contribui para reduzir o nível de ambiguidade da LN, colaborando para a inserção do usuário no ciclo de vida do conhecimento (Ross, 2009; Van Grondelle et al., 2010). Além deste método, algumas abordagens automatizadas que utilizam o processamento da linguagem natural (PLN) também contribuem para tal redução (Jurafsky & Martin, 2000; Manning & Schütze, 2000; Bird et al., 2009).

2.2 Máquina de estados finitos

A MEF é utilizada como um método de notação para modelagem de processo de negócio (Van der Aalst, 2011). Trata-se de um AFD composto por uma quintupla $M = (Q, \Sigma, \delta, q_0, F)$, onde: Q é o conjunto de estados; Σ é o conjunto de entradas ou ações; δ é o conjunto de transições; q_0 é o estado inicial; e F é o estado final (Jurafsky & Martin, 2000).

Além da sua representação formal, a MEF pode ser simbolizada basicamente por um grafo ou por uma tabela de transição de estados. No primeiro caso, o autômato é representado por um grafo $G(V, A)$, em que V é um conjunto não vazio e finito de vértices (nós) e A é um conjunto de pares ordenados de V , chamados arestas (arcos). O tipo de grafo que representa uma MEF é o grafo direcionado, pois as arestas têm uma direção associada. Tal direção é indicada por uma seta na representação do grafo. No segundo caso, o autômato pode ser representado por uma tabela de transição de estados, composta pelos estados, dispostos nas linhas, e pelas entradas (ações) dispostas nas colunas (Jurafsky & Martin, 2000; Hopcroft et al., 2003; Menezes, 2008b).

2.3 Notação combinada

Algumas linguagens são, potencialmente, mais ambíguas que outras. Num extremo, as notações formais de modelagem de processo, como a MEF e a rede de Petri, têm uma semântica bem definida e, portanto, apresentam uma tendência de menor

ambiguidade. No outro extremo, a LN se apresenta com muita ambiguidade inerente. Acredita-se que a substituição da LN por uma notação formal, poderia contribuir para uma redução considerável do nível de ambiguidade na especificação de requisitos de sistemas. Contudo, esta substituição poderia afetar, de forma nociva, a compreensibilidade. Portanto, sugere-se que a melhor abordagem seja a combinação da LN com notações formais. Assim, as vantagens de ambas as linguagens – natural e formal – poderiam ser preservadas. Com a aplicação deste método de notação combinado, que se caracteriza pela restrição da LN com o uso de uma notação formal de modelagem de processo, é provável que haja um equilíbrio entre a compreensibilidade da LN e o potencial de redução da ambiguidade pela notação formal (Davis et al., 1993).

O que se propõe para a NC é a descrição do processo em LN com anotação manual de papéis semânticos. O conjunto de papéis semânticos proposto foi constituído pelos rótulos previstos na notação formal MEF (estado_inicial, estado, ação) e pelo delimitador “fim”, utilizado para indicar o término da abrangência do rótulo precedente. Os rótulos foram atribuídos aos sintagmas que caracterizam, de fato, o processo (Figura 1). O intuito foi verificar se a semântica executável desta notação formal

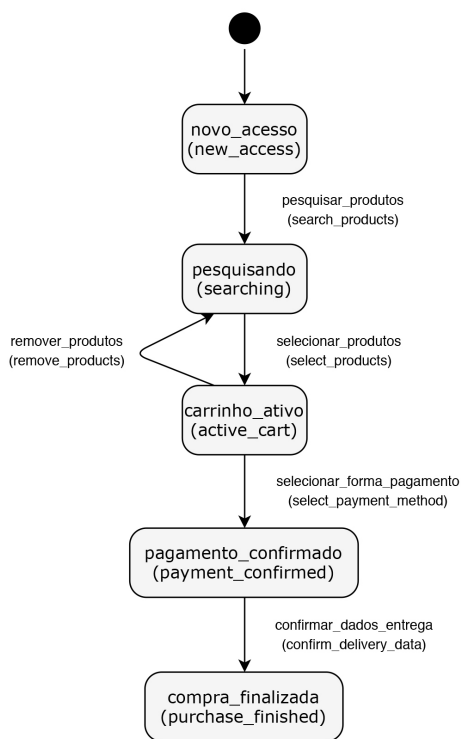


Figura 1. Protótipo de baixa fidelidade do processo de compra numa loja virtual: representação em grafo de uma Máquina de Estados Finitos - estudo experimental.

– metamodelo – possibilita a recuperação de informação não ambígua (Oren et al., 2006; Van der Aalst, 2011; Lautenbacher et al., 2008). Deste modo, instâncias das classes estado e ação podem ser identificadas em trechos do texto. Os rótulos foram atribuídos no editor de texto LibreOffice Writer (Versão: 4.2.4.2) com o comando Inserir > Anotação ou pelo atalho Ctrl + Alt + C (LibreOffice, 2013).

3 Leitura baseada em perspectivas

Trata-se de uma técnica de leitura utilizada na inspeção de sistemas, cujo objetivo é examinar a descrição de um software pela perspectiva das partes interessadas a fim de identificar falhas. A leitura baseada em perspectivas (LBP) é sistemática, focada, personalizável e passível de melhoria a partir da experiência de aplicação e treinamento. Ela pode servir para a inspeção de diversos documentos, incluindo a especificação de requisitos, modelos de projeto e código. Foi elaborada para reduzir a influência humana nos resultados de inspeção. Pesquisas experimentais sugerem que a técnica LBP é significativamente mais eficaz em relação às técnicas de leitura tradicionais como a leitura baseada em *checklist* e a leitura *ad hoc* (Lahtinen, 2012).

A ideia principal da técnica LBP é inspecionar um documento a partir das diferentes perspectivas dos revisores. Na LBP as perspectivas derivam das partes interessadas no documento, ou seja, as pessoas mais relevantes durante o ciclo de vida do projeto. A razão para a adoção de perspectivas distintas é a probabilidade de que um documento seja de maior qualidade quando as partes interessadas não detectam nenhuma falha nele. Perspectivas típicas para a inspeção de um documento de especificação de requisitos poderiam incluir a perspectiva do projetista, testador e/ou usuário. Pelo fato de cada leitor ser responsável por apenas um foco de visão do documento, qualquer erro em potencial é analisado de forma mais rigorosa. Quando é escolhida uma perspectiva única, os revisores tendem a utilizar seu conhecimento específico de forma mais natural e, portanto, mais efetiva. Por outro lado, ao avaliar múltiplas perspectivas de forma simultânea, os revisores são atribuídos à perspectiva mais apropriada (Lahtinen, 2012). Para cada uma dessas perspectivas há um formulário específico com uma série de questões relacionadas que orientam a leitura. Assim, ao utilizar perspectivas distintas haveria uma melhor compreensão das falhas, uma vez que aspectos relativos a cada perspectiva poderiam ser identificados de forma mais eficaz (Basili et al., 1996; Sørungård, 1997; Shull et al., 2000; Sommerville, 2007a).

As falhas identificáveis pela LBP são as seguintes: omissão; informação ambígua; fato incorreto; informação alheia; e diversas (Sørungård, 1997). Trata-se de uma técnica personalizável por permitir que os revisores – desenvolvedores, testadores e usuários – possam adequá-la ao propósito desejado (Shull et al., 2000). Porém, no escopo do presente trabalho, a única

perspectiva utilizada foi a de teste. A chamada leitura baseada em teste (LBT) é focada nos aspectos relativos à testabilidade do sistema, a partir da informação que é apresentada na especificação de requisitos (Basili et al., 1996). Um cenário LBP é um documento de instruções com poucas páginas. Para cada perspectiva um ou mais cenários são escritos para que os revisores possam desempenhar ações específicas e reproduzíveis, além de responder um conjunto de questões para identificar falhas (Lahtinen, 2012). Um cenário consiste de três partes: introdução, instruções e perguntas.

4 Metodologia

4.1 População-alvo

Para a realização desta pesquisa experimental, a população-alvo analisada foi a de 43 estudantes universitários de graduação em Tecnologia em Análise e Desenvolvimento de Sistemas/Sistemas de Informação do Instituto Federal de Educação, Ciência e Tecnologia Fluminense (IFF) do ano letivo de 2013/2014. Os referidos cursos estão inseridos na área do conhecimento designada por “Ciência da Computação”, conforme classificado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

4.2 Variável dependente

A variável dependente escolhida foi o nível de ambiguidade da especificação dos requisitos funcionais do sistema. A escolha desta variável, que no presente estudo foi designada simplesmente por *nível de ambiguidade*, deve-se a diversos fatores. Dentre eles é possível destacar os seguintes: (a) a especificação não-ambígua é um dos critérios de qualidade na descrição dos requisitos previstos no processo de desenvolvimento de software/sistemas (Davis et al., 1993). Portanto, recomenda-se que a especificação de requisitos seja não ambígua (IEEE, 1998); (b) disponibilidade de uma técnica específica de leitura da especificação de requisitos em LN com o propósito de identificar falhas, designada por LBP (Basili, 2013; Sørungård, 1997; Lanubile et al., 1998); (c) a ambiguidade da especificação de requisitos é reportada amplamente na literatura científica como um problema no processo de desenvolvimento de software/sistema (Walia & Carver, 2009; Sutcliffe et al., 1999; Card, 1998); (d) a ambiguidade é elencada como um tipo de falha passível de identificação a partir da utilização da técnica LBP (Sørungård, 1997); (e) após a identificação do requisito ambíguo é possível mensurar o nível de ambiguidade a partir de uma métrica de qualidade aplicável à especificação de requisitos de software (Bokhari & Siddiqui, 2011).

Como a recomendação é de que a especificação de requisitos de software seja não-ambígua, a métrica de qualidade reportada na literatura científica que demonstra adequação para tal propósito é:

$$Q_a = n_{ui} / n_r \quad (1)$$

em que n_{ui} é o número de requisitos para os quais todos os revisores (testadores) apresentaram interpretações idênticas e n_r é o número total de requisitos.

A unidade de requisito foi definida pela transição da MEF pelo fato dela conter uma tríade estado-ação-estado testável. Assim, foi possível obter o percentual de requisitos interpretados de maneira única por todos os testadores. Os valores de referência da Equação 1 são apresentados na escala intervalar Q_a que varia de 0 (zero) até 1 (um), onde os valores próximos de 0 (zero) representam requisitos não ambíguos e os valores próximos de 1 (um) representam requisitos ambíguos.

4.3 Variável independente

A variável independente utilizada no presente estudo foi o método de notação de modelagem de processo. Diferentes métodos foram empregados durante a descrição da especificação dos requisitos funcionais do sistema. Tal variável foi designada simplesmente pelo termo *método de notação* e apresentou três níveis: LN, MEF e NC. Por se tratar de uma linguagem não-artificial, a LN foi utilizada como primeiro nível da variável independente. As demais notações, de caráter artificial, foram atribuídas aos níveis MEF e NC.

MEF é a notação formal de modelagem de processo mais básica (Jurafsky & Martin, 2000). O terceiro nível do método de notação, designado por NC, foi composto por uma combinação de notações. Tal combinação caracteriza-se pela atribuição de rótulos da MEF à descrição em LN. Esta anotação manual de papéis semânticos foi destinada aos sintagmas da LN considerados relevantes pelos seguintes aspectos: (1) caracterizarem os requisitos do processo; e (2) serem passíveis de categorização pelos rótulos da quintupla M que compõem a notação MEF.

A escolha da referida variável independente foi motivada pelos seguintes fatores: (a) acredita-se que a substituição da LN por uma notação formal como a MEF reduza, de forma considerável, o nível de ambiguidade na especificação de requisitos de software. Como tal efeito ocorre às expensas da compreensibilidade, recomenda-se uma combinação da LN com linguagens mais formais como a MEF. Assim, as vantagens da LN e formal são preservadas (Davis et al., 1993); (b) sugere-se que, na LN, haja uma ambiguidade inerente e que a utilização da linguagem MEF apresente como efeito a redução do nível intrínseco de ambiguidade pelo fato de demonstrar uma semântica bem definida (Davis et al., 1993; IEEE, 1998); (c) durante a fase de especificação de requisitos há dois erros comuns: aqueles causados por não saber quais são os requisitos e aqueles causados por não saber como especificá-los de forma adequada (Davis et al., 1993; Sommerville, 2007b). Assim, a proposta do método de notação NC, é demonstrar a viabilidade de especificar requisitos de forma adequada quanto ao nível de ambiguidade.

4.4 Estudo-piloto

Um estudo-piloto populacional de caráter simples-cego foi conduzido com o intuito de ajustar o delineamento experimental. A motivação primária desta abordagem foi adquirir experiência com a aplicação da técnica LBT (tópico 4.8) e fornecer subsídios para nortear a escolha da técnica de inferência estatística mais apropriada no estudo experimental. Neste estudo preliminar analisou-se o pressuposto da normalidade com relação ao nível de ambiguidade na descrição textual do processo em LN. Essa descrição ocorreu durante a especificação de requisitos funcionais de um SIE hipotético. A partir de um protótipo de baixa fidelidade com alto nível de abstração, sujeitos foram motivados a descrever o processo de compra de uma loja virtual (Silva et al., 2013). As implicações do resultado deste estudo preparatório foram de grande relevância para o presente estudo experimental.

A ferramenta utilizada para a análise estatística foi o RStudio: um *Integrated Development Environment* (IDE). Trata-se de um software livre multiplataforma destinado à melhoria da produtividade, com uma interface de usuário apropriada para a utilização da linguagem R (Bryer & Speerschneider, 2012; R Core Team, 2013).

Naquela ocasião, ao especificar os requisitos, o sujeito não precisou se restringir a descrição das funcionalidades previstas no grafo. Este serviu apenas como ponto de partida para que os sujeitos pudessem expressar as funcionalidades do que eles entendem por um sistema de compra numa loja virtual. Após a coleta dos dados de cada um dos sujeitos da amostra, ocorreu a análise das sentenças pela técnica LBT com a utilização da métrica de qualidade Q_a , outrora especificada, para aferir o nível de ambiguidade dos requisitos.

Por tratar-se de um estudo populacional, não houve divisão por grupos no primeiro momento. Além disso, como todo estudo simples-cego, o sujeito não teve informação acerca do fenômeno observado durante o experimento – i.e. nível de ambiguidade. Apenas os pesquisadores detiveram tal informação (Silva et al., 2013).

4.5 Procedimento do estudo experimental

Foi conduzido um estudo experimental simples-cego controlado e aleatorizado. A variável independente, método de notação, foi designada como fator único da análise de variância (ANOVA). Desta forma, a amostra obtida pela população-alvo foi dividida em 3 (três) grupos de sujeitos, cada qual representando um nível do fator único: Grupo Controle LN (GC); Grupo Experimental MEF (GEMEF); e Grupo Experimental NC (GENC). Ao GC foi solicitada a descrição em LN do processo de compra de uma loja virtual, do mesmo modo que foi conduzido no estudo-piloto. No GEMEF foi solicitada a descrição do mesmo processo, porém utilizou-se a notação MEF para a especificação dos requisitos funcionais. No GENC, por sua vez, a descrição do referido

processo foi realizada também em LN, no entanto os sujeitos deste grupo atribuíram rótulos da MEF (estado_inicial, estado, ação) à descrição em LN. O resumo do procedimento experimental adotado na avaliação dos requisitos funcionais do sistema está representado na Tabela 1.

Como apenas 1 (uma) variável independente foi escolhida para a realização do estudo, aplicou-se a ANOVA de um único fator. A variável dependente, por sua vez, foi o nível de ambiguidade na especificação dos requisitos.

O protótipo de baixa fidelidade utilizado no estudo experimental (Figura 1) caracteriza-se por um nível de abstração mais reduzido em relação ao protótipo utilizado no estudo-piloto (Silva et al., 2013).

É possível representar as transições δ pela própria notação formal descrita na Tabela 2 ou pela tabela de transição de estados apresentada na Tabela 3. As transições expressas na Tabela 2 foram os 5 requisitos avaliados pelos testadores. Estes requisitos expressam o protótipo de forma adequada. A Tabela 3 ilustra a tarefa a ser desempenhada pelos sujeitos do experimento que foram designados ao GEMEF. Ao adotá-la é possível observar um exemplo desejável de descrição em MEF do processo de compra de uma loja virtual, como representado pelo protótipo de baixa fidelidade exposto na Figura 1.

A Figura 2 ilustra a aplicação da NC utilizada como parâmetro de avaliação do grupo experimental GENC. Ao adotá-la é possível observar um exemplo

Tabela 1. Procedimento experimental adotado na avaliação da especificação dos requisitos funcionais do sistema.

Procedimento Experimental			
GC	Grupo	Método de Notação	Notação a ser avaliada
GEMEF	Controle	único	LN
GEMEF	Experimental	único	MEF
GENC	Experimental	combinado	LN + MEF

Tabela 2. Requisitos do protótipo de baixa fidelidade representados pelas transições da notação MEF – estudo experimental.

Nº	Transições
1	$\delta(\text{novo_acesso, pesquisar_produtos}) = \text{pesquisando}$
2	$\delta(\text{pesquisando, selecionar_produtos}) = \text{carrinho_ativo}$
3	$\delta(\text{carrinho_ativo, remover_produtos}) = \text{pesquisando}$
4	$\delta(\text{carrinho_ativo, selecionar_forma_pagamento}) = \text{pagamento_confirmado}$
5	$\delta(\text{pagamento_confirmado, confirmar_dados_entrega}) = \text{compra_finalizada}$

Tabela 3. Tabela de transição de estados: representação da notação MEF que caracteriza o processo de compra numa loja virtual – estudo experimental.

		Ações (Σ)				
δ		pesquisar_ produtos	selecionar_ produtos	remover_ produtos	selecionar_ forma_ pagamento	confirmar_ dados_ entrega
Estados (Q)	novo_acesso	pesquisando	\emptyset	\emptyset	\emptyset	\emptyset
	pesquisando	\emptyset	carrinho_ativo	\emptyset	\emptyset	\emptyset
	carrinho_ativo	\emptyset	\emptyset	pesquisando	pagamento_ confirmado	\emptyset
	pagamento_ confirmado	\emptyset	\emptyset	\emptyset	\emptyset	compra_ finalizada
	compra_finalizada	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset

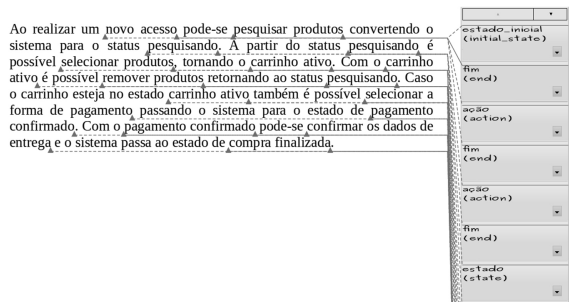


Figura 2. Descrição do processo de compra de uma loja virtual em notação combinada, utilizada como parâmetro de avaliação do grupo experimental GENC.

desejável de descrição em LN do processo de compra numa loja virtual, como representado pelo protótipo de baixa fidelidade exposto na Figura 1. Além da LN é possível observar também os rótulos previstos na notação formal MEF sendo atribuídos aos sintagmas que caracterizam, de fato, o processo de negócio. O rótulo “fim” é apenas um ajuste para indicar o término da abrangência do rótulo precedente.

4.6 Delineamento experimental

A aplicação da ANOVA depende do cumprimento de três pressupostos: normalidade, independência e homocedasticidade (Larson & Farber, 2010). O delineamento completamente casualizado (DCC) foi adotado para a execução da rotina prevista na ANOVA de fator único. A escolha deste tipo de delineamento deve-se ao objetivo de caráter comparativo proposto no presente estudo. É provável que outros fatores também afetem o nível de ambiguidade, porém o que se pretende é saber se o fator primário *método de notação*, definido *a priori*, afeta o nível de ambiguidade de forma significativa, apesar da eventual existência de outros fatores intervenientes (NIST, 2013).

A aplicação do DCC contribui sobremaneira para o cumprimento do pressuposto da independência. Tal fato deve-se à aleatoriedade da atribuição dos níveis do fator primário às unidades experimentais. Portanto, o delineamento é completamente casualizado ou aleatorizado (Bailey, 2008b; NIST, 2013). Deste modo, todos os níveis do fator primário têm a mesma chance de serem aplicados em qualquer unidade experimental – sujeitos do experimento (Bailey, 2008a, b).

Finalmente, para que não haja qualquer restrição inicial acerca da correta aplicação da ANOVA de fator único, o seguinte procedimento foi adotado: após a coleta dos dados foi verificado o cumprimento do pressuposto da homoscedasticidade. Antes de comparar as médias é preciso checar a suposição de que a homogeneidade de variâncias é razoável. Com este intuito, o teste de Levene é passível de aplicação (NIST, 2013; Bisquerra et al., 2007).

O teste *post-hoc* de Tukey foi utilizado para identificar se há diferenças estatisticamente significativas intergrupo e entre quais grupos estas diferenças se apresentam. Dentre os testes *post-hoc* mais comuns, o teste de Tukey foi escolhido pelo fato de preservar o nível de significância predefinido de 5%, apesar do aumento do número de médias a serem analisadas. Ou seja, o nível de significância definido *a priori* é mantido, apesar do incremento do número de grupos a serem analisados.

4.7 Atribuições do avaliador

Na coleta de dados do estudo-piloto foi solicitada a descrição do protótipo apenas em LN. Do mesmo modo, ocorreu a descrição do protótipo pelo GC do estudo experimental. No entanto, aos sujeitos do estudo experimental designados aos grupos experimentais GENC e GEMEF, foi concedida uma breve orientação sobre a tarefa a ser desempenhada. Além disso, durante todo o procedimento, os grupos experimentais tiveram à disposição um material de referência com a descrição de outro processo, utilizando o método de notação específico do grupo para o qual foi designado. Nenhum grupo teve informação acerca da tarefa desempenhada por outro grupo e não houve limitação com relação ao tempo de execução da tarefa.

4.8 Atribuições do leitor-testador

A partir da aplicação da técnica LBT e com base nos requisitos especificados pelos sujeitos, o leitor-testador adotou o seguinte procedimento: (1) em caso de falhas, estas foram classificadas, seguindo o cenário específico para o leitor-testador; (2) toma-se como referência o protótipo de baixa fidelidade e seus requisitos (Tabela 2); (3) as falhas identificadas e classificadas durante a leitura foram reportadas em formulário apropriado.

Conforme mencionado anteriormente, no tópico 4.2, que trata da variável dependente, a unidade de requisito foi definida pela transição da MEF, pelo fato dela conter uma tríade estado-ação-estado testável. Se, por algum motivo, o requisito especificado pelo sujeito não contemplar o conteúdo expresso na tríade, a falha é contabilizada. Na avaliação do GENC, apenas os sintagmas rotulados foram avaliados pelos testadores.

As falhas identificáveis foram: omissão; informação ambígua; fato incorreto; informação alheia; e diversas. Aos testadores foi disponibilizado um formulário com a descrição do procedimento a ser adotado para detectar essas falhas na especificação dos requisitos (Lahtinen, 2012).

5 Hipóteses da pesquisa

5.1 Hipótese básica

A partir do problema abordado de forma explícita no tópico 1, foi possível delinear o par de hipóteses de caráter geral acerca do estudo em seu contexto mais amplo. $H_{0(i)}$: *O método de notação, utilizado na modelagem de processos, não afeta o nível de ambiguidade do processo descrito na especificação de requisitos funcionais.* $H_{1(i)}$: *O método de notação, utilizado na modelagem de processos, afeta o nível de ambiguidade do processo descrito na especificação de requisitos funcionais.*

5.2 Hipóteses secundárias

Para o cumprimento do objetivo específico previsto no tópico 1 (item i), foi conduzido um estudo-piloto que teve como ponto de partida as seguintes hipóteses: (i) $H_{0(p)}$: *A amostra não é proveniente de uma população-alvo normalmente distribuída quanto ao nível de ambiguidade na especificação de requisitos funcionais em LN.* $H_{1(p)}$: *A amostra é proveniente de uma população-alvo normalmente distribuída quanto ao nível de ambiguidade na especificação de requisitos funcionais em LN.* Após a proposição do método combinado de notação para a modelagem de processos, foi conduzido um estudo experimental para o cumprimento dos objetivos específicos iii e iv previstos no tópico 1. Este estudo experimental teve como ponto de partida as seguintes hipóteses: (ii) $H_{0(ii)}$: *Não há diferença intergrupo no nível de ambiguidade do processo descrito na especificação de requisitos funcionais.* $H_{1(ii)}$: *Há diferença intergrupo no nível de ambiguidade do processo descrito na especificação de requisitos funcionais.*

6 Resultados

6.1 Análise paramétrica de variância

Na ANOVA de fator único e DCC, a hipótese básica $H_{0(i)}$ não foi rejeitada ao considerar o nível de significância de 5%, pois $p > \alpha$ ($0,07 > 0,05$) e $F < F$ crítico ($2,89 < 3,24$). Este resultado sugere que o método de notação, utilizado na modelagem de processos, não afeta o nível de ambiguidade do processo descrito na especificação de requisitos funcionais ($p > 0,05$). Na Figura 3 o nível médio de ambiguidade e o erro

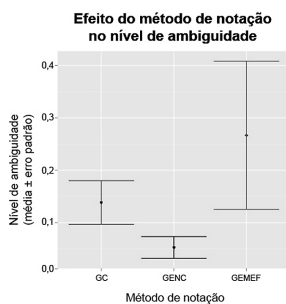


Figura 3. Efeito do método de notação de modelagem de processo no nível de ambiguidade da especificação dos requisitos funcionais de sistemas de informação empresariais.

padrão são apresentados: GC ($0,14 \pm 0,04$), GENC ($0,05 \pm 0,02$) e GEMEF ($0,27 \pm 0,14$). A comparação de médias sugere que: a NC é o melhor método para reduzir o nível de ambiguidade da especificação de requisitos; a especificação de requisitos expressa em MEF gera o maior o nível de ambiguidade dentre as notações avaliadas. No entanto, estas tendências não são estatisticamente significativas ($p > 0,05$). Na Tabela 4 foi apresentado o resumo dos dados estatísticos obtidos.

6.2 Análise não-paramétrica de variância

O teste de Kruskal-Wallis apontou resultados análogos à ANOVA de fator único. A hipótese básica $H_{0(i)}$ não foi rejeitada ao considerar o nível de significância de 5%, pois $p > \alpha$ ($0,20 > 0,05$). Assim, este resultado também sugere que o método de notação, utilizado na modelagem de processos, não afeta o nível de ambiguidade do processo descrito na especificação de requisitos funcionais ($p > 0,05$).

6.3 Homogeneidade de variâncias

Pelo teste de Levene padrão, que utiliza a média como parâmetro, a hipótese de homocedasticidade foi rejeitada, pois $p < \alpha$ ($0,00006 < 0,05$). Por outro lado, no teste de Levene calculado pela mediana (teste Brown–Forsythe), a hipótese de homocedasticidade não foi rejeitada, pois $p > \alpha$ ($0,056 > 0,05$).

6.4 Normalidade

No teste de Shapiro-Wilk a hipótese nula afirma que a amostra é proveniente de uma população normalmente distribuída. Por outro lado, a hipótese nula $H_{0(p)}$ (tópico 5.2, item i) do presente estudo é contrária, pois afirma que a amostra não é proveniente de uma população normalmente distribuída.

O resultado do teste de aderência para cada um dos grupos experimentais está expresso na Tabela 5. A hipótese nula de Shapiro-Wilk foi rejeitada ao nível de significância de 5%, porque em todos os grupos experimentais o W calculado é menor que o W crítico tabelado. Este fato implicou em não rejeitar a hipótese nula $H_{0(p)}$ do presente estudo, sugerindo que a nenhum dos grupos é proveniente de uma população normalmente distribuída.

Tabela 4. Resumo dos dados estatísticos observados.

Grupo	n	Média	Mediana	Máximo	Desvio-padrão (σ)	Variância (σ^2)
GC	13	0,14	0,20	0,40	0,15	0,02
GENC	19	0,05	0,00	0,20	0,09	0,01
GEMEF	9	0,27	0,00	1,00	0,42	0,18

Tabela 5. Resultado do teste de aderência de Shapiro-Wilk em cada um dos grupos experimentais.

Grupo	n	W	W*	P	p*
GC	13	0,79	W < 0,866	0,0052260	p < 0,05
GENC	19	0,55	W < 0,901	0,0000016	p < 0,05
GEMEF	9	0,66	W < 0,829	0,0004289	p < 0,05

p* = comparação do p calculado com o nível de significância ($\alpha = 0,05$); W* = comparação do W com W crítico; n = número de indivíduos na amostra; W = estatística do teste; P = probabilidade calculada.

Tabela 6. Teste *post-hoc* de Tukey: comparação pareada do nível médio de ambiguidade.

Diferença intergrupo	p	p*	Erro Tipo II (β)	Poder do Teste
GENC-GC	0,533	p > 0,05	0,59	0,41
GEMEF-GC	0,383	p > 0,05	0,58	0,42
GEMEF-GENC	0,056	p > 0,05	0,65	0,35

p* = comparação do p calculado com o nível de significância ($\alpha = 0,05$).

6.5 Análise residual

A análise quantil-quantil revelou, a não aderência do perfil do nível de ambiguidade dos grupos GENC e GEMEF com a curva de Gauss. A não aderência também foi observada no GC. A fuga de normalidade pode ser atribuída ao fato dos dados que compõem a métrica do nível de ambiguidade serem de caráter discreto.

6.6 Análise *post-hoc*

O teste *post-hoc* de Tukey sugere que não há diferença estatisticamente significativa intergrupo quanto ao nível de ambiguidade do processo descrito na especificação de requisitos funcionais ($p > 0,05$). Assim, o resultado da comparação pareada do nível médio de ambiguidade sugere que a hipótese nula $H_{0(ii)}$ (tópico 5.2, item ii) não foi rejeitada.

Ao adotar o nível de significância de 5%, não houve nenhuma diferença estatisticamente significativa entre os pares analisados (Tabela 6). O fato da hipótese nula não ter sido rejeitada ressaltou a importância de se avaliar o erro tipo II (β) – probabilidade de não rejeitar uma falsa hipótese nula. Assim, estimou-se o poder do teste estatístico definido por $1-\beta$, ou seja, a probabilidade de encontrar uma diferença intergrupo que, de fato, proceda.

O perfil da distribuição do nível de ambiguidade intergrupo foi representado na Figura 4 pelo polígono de frequência relativa. Tomando como referência o eixo das abscissas, nota-se que houve uma sobreposição da concentração dos resíduos empíricos. Além disso, também é notória a assimetria positiva da distribuição. Em termos gerais, o grau de assimetria observado foi de 2.61.

6.7 Escala de Likert

Quatro questões críticas foram observadas pelos testadores durante a análise da amostra. Estas questões referentes à aplicação da técnica LBT foram avaliadas

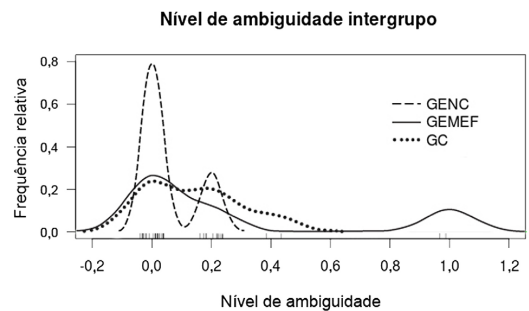


Figura 4. Distribuição do nível de ambiguidade intergrupo.

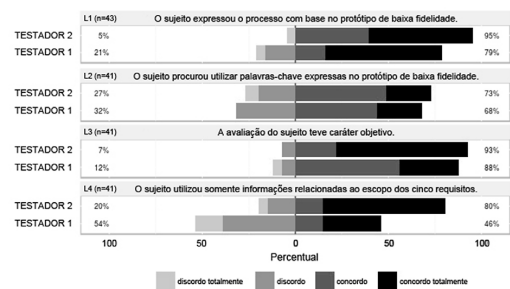


Figura 5. Escala de Likert com questões críticas experimentadas na utilização da técnica LBT durante a rotina de análise dos dados do estudo experimental.

de forma qualitativa pelo uso da escala de Likert da Figura 5. De forma geral, é possível notar o otimismo do testador 2 em relação ao testador 1 no que se refere as questões críticas. A primeira questão serviu como critério para descartar sujeitos da amostra quando ambos os testadores concordaram com o fato do sujeito não ter descrito o processo de compra com base no protótipo de baixa fidelidade. Deste modo, 4,65% dos 43 sujeitos foram excluídos das análises subsequentes. Assim, a avaliação dos demais itens de Likert foi realizada com os 41 sujeitos restantes. Com relação ao segundo item apresentado na Figura 5,

observa-se que mesmo dentre os sujeitos válidos, houve um consenso entre os testadores de que 17,07% da amostra não utilizou palavras-chave, fato que afetou a testabilidade dos cinco requisitos previstos na Tabela 2. Outro consenso entre os testadores foi que 85,37% das avaliações tiveram um caráter objetivo. Neste quesito, o nível de subjetividade foi de 4,88%. Por fim, em 19,51% dos casos foi possível notar a presença de informações fora do escopo dos cinco requisitos avaliados.

7 Discussão

No estudo-piloto, 83,87% dos sujeitos da amostra foram abordados pelo mesmo avaliador, minimizando possíveis discrepâncias na abordagem. Apesar do nível de subjetividade da avaliação ter sido relativamente baixo, não houve consenso entre os testadores em 29,17% dos casos para este item. A correlação de postos de Spearman foi de 0,45 neste quesito, sugerindo uma baixa relação entre a opinião dos testadores ao avaliar a subjetividade. A dificuldade em avaliá-la pode ter sido afetada pelo fato de 58,33% dos sujeitos apresentarem informação fora do escopo dos cinco requisitos passíveis de avaliação (Silva et al., 2013).

Ao analisar o cumprimento dos pressupostos da ANOVA reportados no tópico 4.6, com base nos resultados do estudo-piloto e do estudo experimental é possível considerar os seguintes aspectos: (1) *Normalidade*: o teste de aderência de Shapiro-Wilk e a análise quantil-quantil para a análise residual demonstraram que a distribuição do valor da variável nível de ambiguidade não apresentou perfil normal em nenhum dos grupos analisados ($p < 0,05$). A não-normalidade da especificação dos requisitos funcionais em LN, observada no estudo-piloto, confirmou-se no grupo controle do estudo experimental e também foi observada nos grupos experimentais GENC e GEMEF; (2) *Independência*: a amostragem de dados foi aleatória para assegurar a independência das observações. A aleatorização prevista no processo de DCC assegurou que a ocorrência de uma observação não afetou a probabilidade de ocorrência de outra observação; (3) *Homocedasticidade*: o teste de Levene calculado pela mediana (teste de Brown-Forsythe) demonstrou que a hipótese de homocedasticidade não foi rejeitada ($p > 0,05$). Este teste foi adotado por apresentar boa robustez em dados com distribuição não normal. O teste de Levene padrão, que utiliza a média como parâmetro, rejeitou a hipótese de homocedasticidade. No entanto, ele não foi considerado, pois é mais indicado para distribuições que apresentam um perfil normal – situação que não ocorreu (NIST, 2013). Assim, não houve diferença intergrupo com relação à variância. Portanto, as variâncias são homogêneas, fato que reduz o risco de que as variações atribuídas às diferenças individuais sejam maiores do que aquelas relativas a aplicação

do fator propriamente dito. Essa homogeneidade é fundamental para favorecer a validade de conclusão, mas compromete a validade externa, pois representa uma ameaça à heterogeneidade aleatória dos sujeitos.

Com base nos aspectos mencionados é possível inferir que apenas o pressuposto da normalidade não foi cumprido. Como os pressupostos para a aplicação da ANOVA (análise paramétrica) não foram cumpridos de forma integral, a análise não-paramétrica seria a mais apropriada. No entanto, como o teste de aderência de Shapiro-Wilk é muito sensível à desvios de normalidade, a rejeição à hipótese de normalidade é muito provável. Assim, optou-se por conduzir as análises paramétrica (ANOVA de fator único e DCC) e não paramétrica (teste de Kruskal-Wallis). Os resultados obtidos por ambas foram similares.

Tanto no estudo-piloto quanto no estudo experimental houve restrições quanto à disponibilidade de recursos para a realização do experimento, de maneira simultânea, com todos os sujeitos. Situação que motivou a aplicação do experimento em ocasiões distintas. Contudo, no caso do estudo experimental, todos os sujeitos da amostra foram abordados pelo mesmo avaliador, minimizando possíveis discrepâncias na abordagem. Apesar do nível de subjetividade da avaliação ter sido baixo, não houve consenso entre os testadores em 9,76% dos casos para este item. A correlação de postos de Spearman foi de 0,40 neste quesito, sugerindo uma baixa relação entre a opinião dos testadores ao avaliar a subjetividade. Assim como no estudo-piloto, a dificuldade em avaliá-la pode ter sido afetada pelo fato de 19,51% dos sujeitos apresentarem informação fora do escopo dos cinco requisitos passíveis de avaliação.

No estudo-piloto, o principal fator que pode ter afetado, de forma negativa, todos os itens de Likert avaliados foi a complacência nas orientações aos sujeitos acerca de como a tarefa deveria ser desempenhada. Naquela ocasião, como foi utilizado um protótipo de baixa fidelidade com alto nível de abstração, foi permitido aos sujeitos que descrevessem tarefas do processo de compra, que poderiam estar implícitas no modelo (Silva et al., 2013). No entanto, no estudo experimental os sujeitos foram orientados à descrever estritamente o que estava representado no protótipo de baixa fidelidade. Além disso, o último protótipo utilizado caracterizou-se por um nível de abstração mais baixo, a fim de representar o processo de compra de forma mais realística. Assim, os dados sugerem que o efeito combinado da redução do nível de complacência nas orientações com a redução do nível de abstração do protótipo podem ter afetado, de forma positiva, todos os itens de Likert avaliados.

A técnica LBT demonstrou uma aplicabilidade favorável enquanto ferramenta específica para a análise proposta. No entanto, a utilização da métrica do nível de ambiguidade Q_a associada à técnica LBT

subestimou o nível de ambiguidade, situação que foi evidenciada pela assimetria positiva da distribuição, cujo valor observado foi de 2,61 (Figura 4). Como o numerador de $Q_a(n_{ui})$ só é contabilizado quando há concordância entre os testadores em relação à ambiguidade do requisito, a opinião do testador mais otimista tende a prevalecer. Além disso, a probabilidade de atribuir ambiguidade a um requisito é muito pequena em termos matemáticos. Há seis rótulos possíveis na LBT, caracterizando o seguinte espaço amostral: LBT = {omissão, informação ambígua, fato incorreto, informação alheia, outras falhas, nenhuma falha}. Entre os testadores pode haver T = {concordância, discordância, não-consenso}, totalizando outros três rótulos possíveis. A ambiguidade só é contabilizada em n_{ui} , se o rótulo da LBT for “informação ambígua” e houver “concordância” entre os dois testadores. Sendo LBT(informação ambígua) \cong 16,67% e T(concordância) \cong 33,33%, a probabilidade de contabilizar a ambiguidade em n_{ui} é definido por $P(\text{ambiguidade}) = \text{LBT}(\text{informação ambígua}) \cap \text{T}(\text{concordância}) \cong$ 5,56%. Uma alternativa para minimizar este problema seria reduzir o número de rótulos possíveis na LBT. Algo como LBT = {informação ambígua, outras falhas, nenhuma falha}. Assim, a probabilidade de contabilizar a ambiguidade em n_{ui} atingiria aproximadamente 11,11%, forçando a redução da assimetria no perfil da distribuição dos dados para favorecer a realização de análises comparativas sob uma perspectiva mais favorável.

Algumas situações parecem ocorrer em caráter circunstancial (Figura 3). É o caso da baixa ambiguidade relativa da especificação de requisitos em LN, em face ao conceito preestabelecido de que a LN apresenta uma alta ambiguidade. Contraponto também observado na alta ambiguidade relativa da notação formal, em face ao conceito preestabelecido de que a notação formal apresenta uma baixa ambiguidade. Numa análise meramente qualitativa e independente de significância estatística, os resultados deste estudo apontam que a ambiguidade inerente à LN e a ausência de ambiguidade inerente à notação formal são altamente dependentes do uso. É provável que estas propriedades potenciais e intrínsecas dos métodos de notação tendam a se manifestar de acordo com o grau de proficiência e/ou de maturação do usuário.

Assim, deve-se avaliar o efeito do treinamento no nível de ambiguidade. A LN pode ter apresentado um nível de ambiguidade relativamente baixo por ser uma notação mais usual para os sujeitos. A notação formal, ainda que apresente um bom potencial para a redução da ambiguidade, precisa ser exercitada para evidenciar seus efeitos. Deste modo, a avaliação do efeito do fator treinamento no nível de ambiguidade é imprescindível para estimar, de forma mais robusta, o “efeito notação combinada” enquanto método de notação alternativo. A mensuração da variável nível

de ambiguidade é uma tarefa difícil por ser muito sensível aos fatores intervenientes. É necessário que o escopo experimental seja muito bem elaborado e definido para que a fidedignidade e a reprodutibilidade científica das avaliações sejam preservadas. O zelo por estes aspectos torna o ambiente mais assertivo para observações do fenômeno.

Como os sujeitos são estudantes de graduação em ciência da computação e a especificação dos requisitos funcionais pode ser realizada por programadores, testadores, analistas e engenheiros de diversas áreas, não se pode generalizar os resultados deste estudo. Porém, a disponibilidade de recursos humanos restringe sobremaneira a seleção de uma amostra que represente a população-alvo desejável de forma fidedigna.

8 Considerações finais

A partir dos problemas levantados no estudo-piloto, os resultados obtidos apontam que a técnica LBT demonstrou uma aplicabilidade satisfatória para analisar o nível de ambiguidade da especificação dos requisitos funcionais de SIEs. Além disso, os resultados sugerem que a amostra não é proveniente de uma população-alvo normalmente distribuída quanto ao nível de ambiguidade em LN. No estudo experimental, observou-se o efeito do método de notação de modelagem de processo no nível de ambiguidade da especificação de requisitos. Via de regra, no desenvolvimento de sistemas, os requisitos são expressos em LN e é desejável que sua especificação apresente o menor nível possível de ambiguidade. Portanto, esta pesquisa experimental propôs uma combinação da LN, amplamente utilizada para tal propósito, com outro método de notação para induzir a redução do nível de ambiguidade – a máquina de estados finitos.

Os resultados apontam que não houve evidência estatisticamente significativa para afirmar que o método de notação, utilizado na modelagem de processos, afete o nível de ambiguidade. Além disso, não houve diferença estatisticamente significativa intergrupo no nível de ambiguidade. Assim, não se pode afirmar que a combinação da LN com a MEF reduza o nível de ambiguidade da especificação dos requisitos funcionais de SIEs.

Para a realização de estudos posteriores neste campo do conhecimento é importante considerar: o efeito do treinamento no método de notação para induzir ajustes no nível de ambiguidade; a redução do número de rótulos possíveis na LBT a fim de aumentar a probabilidade matemática de consenso (concordância/discordância) entre testadores com relação às falhas. Esta redução poderá contribuir para minimizar a assimetria no perfil da distribuição do nível de ambiguidade, possibilitando a realização de análises comparativas sob uma perspectiva mais

favorável. Tal ação deve contribuir para evitar que o nível de ambiguidade seja subestimado e altamente condicionado à opinião do testador mais otimista. Além disso, é relevante manter reduzido o nível de complacência nas orientações e o nível de abstração do protótipo de baixa fidelidade para melhorar a qualidade da análise.

A anotação semântica da especificação de requisitos em LN com rótulos da SBPM poderia ser considerada em futuros estudos. Quando há relação estreita entre os rótulos da notação formal (metamodelo) e a ontologia de domínio, talvez haja a possibilidade de recuperar informação de forma mais eficaz e menos ambígua, facilitando o fluxo de conhecimento organizacional. Portanto, a análise comparativa e experimental do nível de ambiguidade a partir da anotação semântica baseada em diferentes notações formais de modelagem de processos de negócio seria plausível. Em contrapartida, o grau de proficiência dos sujeitos precisa ser ponderado, pois há notações formais com muito mais rótulos e isso pode dificultar a escolha do rótulo mais apropriado para o sintagma durante a anotação, caso os indivíduos não estejam treinados ou não haja uma ontologia definida.

Referências

- Bailey, R. A. (2008a). Forward look. In R. A. Bailey, *Design of comparative experiments* (Cap. 1, pp. 1-16). Cambridge: Cambridge University Press.
- Bailey, R. A. (2008b). Unstructured experiments. In R. A. Bailey, *Design of comparative experiments* (Cap. 2, pp. 19-41). Cambridge: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511611483.003>.
- Basili, V. R. (2013). *Lab package for the empirical investigation of perspective-based reading*. College Park: University of Maryland. Recuperado em 25 de março de 2013, de <http://www.cs.umd.edu/projects/SoftEng/ESEGE/>
- Basili, V. R., Green, S., Laitenberger, O., Lanubile, F., Shull, F., Sørumgård, S., & Zelkowitz, M. V. (1996). The empirical investigation of perspective-based reading. *Empirical Software Engineering: an International Journal*, 1(2), 133-164. <http://dx.doi.org/10.1007/BF00368702>.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python*. Sebastopol: O'Reilly Media.
- Bisquerra, R., Sarriera, J. C., & Matínez, F. (2007). Provas de homoscedasticidade. In R. Bisquerra, J. C. Sarriera & F. Martínez (Eds.), *Introdução a estatística: enfoque informático com o pacote estatístico SPSS* (Cap. 5, pp. 83-87). Porto Alegre: Bookman.
- Bokhari, M. U., & Siddiqui, S. T. (2011). Metrics for requirements engineering and automated requirements tools. In *Proceedings of the 5th National Conference on Computing for Nation Development (INDIACOM)*. New Delhi: Bharati Vidyapeeth's Institute of Computer Applications and Management.
- Bryer, J., & Speerschneider, K. (2012). *Likert: functions to analyze and visualize likert type items*. R package *version 1.1*. Recuperado em 7 de julho de 2013, de <http://jason.bryer.org/likert>
- Card, D. N. (1998). Learning from our mistakes with defect causal analysis. *IEEE Software*, 15(1), 56-63. <http://dx.doi.org/10.1109/52.646883>.
- Carvalho, R. A., Carvalho, F. L. C., & Manhães, R. S. (2011). Business language driven development: joining business process models to automated tests. In *Proceedings of the VIFIP TC8 International Conference on Enterprise Information Systems*. Aalborg: LNBIP.
- Carvalho, R. A., Manhães, R. S., & Silva, F. L. C. (2010). *Introducing business language driven development*. Ithaca: Cornell University. Relatório técnico. Recuperado em 25 de maio de 2013, de <http://arxiv.org/abs/1011.2238>
- Carvalho, R. A., Silva, F. L. C., & Manhães, R. S. (2012). Advances in enterprise information systems II. In C. Møller & S. S. Chaudhry (Eds.), *Business language driven development: joining business process models to automated tests* (pp. 167-177). Leiden: CRC Press.
- Cohn, M. (2004). *User stories applied for agile software development: an overview* (Cap. 1, pp. 3-15). Boston: Addison-Wesley.
- Coughlan, J., & Macredie, R. D. (2002). Effective communication in requirements elicitation: a comparison of methodologies. *Journal of Requirements Engineering*, 7(2), 47-60. <http://dx.doi.org/10.1007/s007660200004>.
- Coughlan, J., Lycett, M., & Macredie, R. D. (2003). Communication issues in requirements elicitation: a content analysis of stakeholder experiences. *Information and Software Technology*, 45(8), 525-537. [http://dx.doi.org/10.1016/S0950-5849\(03\)00032-6](http://dx.doi.org/10.1016/S0950-5849(03)00032-6).
- Davis, A., Overmyer, S., Jordan, K., Caruso, J., Dandashi, F., Dinh, A., Kincaid, G., Ledebor, G., Reynolds, P., Sitaram, A., Ta, A., & Theofanos, M. (1993). Identifying and measuring quality in a software requirements specification. In *Proceedings of the 1st International Software Metrics Symposium*. New York: IEEE. <http://dx.doi.org/10.1109/METRIC.1993.263792>.
- Hopcroft, J. E., Ullman, J. D., & Motwani, R. (2003). Autômatos finitos. In J. Ullman & J. Hopcroft (Eds.), *Introdução à teoria de autômatos, linguagens e computação* (Cap. 2, pp. 39-88). Rio de Janeiro: Elsevier.
- Institute of Electrical and Electronics Engineers – IEEE. (1998). *IEEE STD 830: IEEE recommended practice for software requirements specifications* (pp. 1-40). Piscataway: IEEE.
- Jørgensen, H. D., Karlsen, D., & Lillehagen, F. (2007). Product based interoperability: approaches and requirements. In *Proceedings of the Conference Coordination of Collaborative Engineering: State of the Art and Future Challenges 5th International Workshop on Challenges in Collaborative Engineering (CCE'07)* (pp. 33-44). Cracow: CCE.
- Jurafsky, D., & Martin, J. H. (2000). *Speech and language processing: an introduction to natural language processing, computational linguistics and speech recognition*. New Jersey: Prentice Hall.
- Kotler, P., Kartajaya, H., & Setiawan, I. (2012). O marketing da missão junto aos consumidores. In P. Kotler, H. Kartajaya & I. Setiawan, *Marketing 3.0: as forças*

- que estão definindo o novo marketing centrado no ser humano (Cap. 3, pp. 57-78). Rio de Janeiro: Elsevier.
- Lahtinen, J. (2012). Application of the perspective-based reading technique in the nuclear I&C context. In VTT Technical Research Centre of Finland (Ed.), *CORSICA Work Report 2011* (VTT Technology, 9). Finland: VTT Technical Research Centre of Finland.
- Laubule, F., Shull, F., & Basili, V. R. (1998). Experimenting with error abstraction in requirements documents. In *Proceedings of 5th International Software Metrics Symposium (METRICS'98)* (pp. 114-121). Bethesda: IEEE Computer Society. <http://dx.doi.org/10.1109/METRIC.1998.731236>.
- Larson, R., & Farber, B. (2010). Testes qui-quadrado e a distribuição F. In R. Larson & B. Farber, *Estatística aplicada* (Cap. 10, pp. 439-483). São Paulo: Pearson Prentice Hall.
- Lautenbacher, F., Bauer, B., & Seitz, C. (2008). Semantic business process modeling: benefits and capability. In *Proceedings of the AAAI 2008 Stanford Spring Symposium: AI Meets Business Rules and Process Management (AIBR)*. California: Stanford University.
- Leffingwell, D. (2011). A brief history of software requirements methods. In D. Leffingwell, *Agile software requirements: lean requirements practices for teams, programs, and the enterprise* (Cap. 1, pp. 3-28). Boston: Addison-Wesley.
- LibreOffice. (2013). Recuperado em 2 de abril de 2013, de <http://www.libreoffice.org/>
- Lopes, E. (2008). Modalidades de gramática. In E. Lopes, *Fundamentos da linguística contemporânea* (Cap. 5, pp. 183-228). São Paulo: Pensamento-Cultrix.
- Manning, C. D., & Schütze, H. (2000). *Foundations of statistical natural language processing*. Cambridge: The MIT Press.
- Menezes, P. B. (2008a). Hierarquia de classes de linguagens e conclusões. In P. B. Menezes, *Linguagens formais e autômatos* (Cap. 9, pp. 195-201). Porto Alegre: Bookman.
- Menezes, P. B. (2008b). Linguagens regulares. In P. B. Menezes, *Linguagens formais e autômatos* (Cap. 3, pp. 41-79). Porto Alegre: Bookman.
- National Institute of Standards and Technology – NIST. SEMATECH. (2013). *e-handbook of statistical methods*. Gaithersburg: NIST. Recuperado em 17 de março de 2013, de <http://www.itl.nist.gov/div898/handbook/>
- Oren, E., Möller, K. H., Scerri, S., Handschuh, S., & Sintek, M. (2006). *What are semantic annotations?* Ireland: DERI Galway. Technical report.
- Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, 122(3), 280-291. <http://dx.doi.org/10.1016/j.cognition.2011.10.004>. PMID:22192697.
- R Core Team. (2013). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Recuperado em 7 de julho de 2013, de <http://www.R-project.org>
- Raffin, F. (2009). Rigor e ambiguidade. In *Pequena introdução à filosofia* (Cap. 15, pp. 181-191). Rio de Janeiro: Editora FGV.
- Ross, R. G. (2009). RuleSpeak® sentence forms: specifying natural-language business rules in english. *Business Rules Journal*, 10(4). Recuperado em 28 de maio de 2013, de <http://www.BRCommunity.com/a2009/b472.html>
- Shull, F., Rus, I., & Basili, V. (2000). How perspective-based reading can improve requirement inspections. *Computer*, 33(7), 73-79. <http://dx.doi.org/10.1109/2.869376>.
- Silva, A. E. C., Carvalho, R. A., & Ferreira, H. S. (2013). Nível de ambiguidade da especificação dos requisitos funcionais de sistemas de informação empresarial: um estudo-piloto. In *Anais do XX Simpósio de Engenharia de Produção (SIMPEP)*. Bauru: Unesp. Recuperado em 4 de junho de 2014, de http://www.simpep.feb.unesp.br/abrir_arquivo_pdf.php?tipo=artigo&evento=8&art=933&cad=17616&opcao=com_id
- Sommerville, I. (2007a). Verificação e validação. In I. Sommerville, *Engenharia de software* (Cap. 22, pp. 341-453). São Paulo: Addison-Wesley.
- Sommerville, I. (2007b). Requisitos de software. In I. Sommerville, *Engenharia de software* (Cap. 6, pp. 79-94). São Paulo: Addison-Wesley.
- Sørungård, S. (1997). *Verification of process conformance in empirical studies of software development* (Tese de doutorado). Trondheim: Department of Computer and Information Science, The Norwegian University of Science and Technology.
- Sutcliffe, G. A., Economou, A., & Markis, P. (1999). Tracing requirements errors to problems in the requirements engineering process. *Requirements Engineering*, 4(3), 134-151. <http://dx.doi.org/10.1007/s007660050024>.
- Turnquist, G. L. (2011). Testing customer stories with behavior driven development. In G. L. Turnquist, *Python testing cookbook* (Cap. 4, pp. 117-163). Birmingham: Packt Publishing.
- Van der Aalst, W. M. P. (2011). Process modeling and analysis. In W. M. P. Aalst, *Process mining: discovery, conformance and enhancement of business processes* (Cap. 2, pp. 29-57). Heidelberg: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-19345-3_2.
- Van Grondelle, J., Heller, R., Van Haandel, E., & Verburg, T. (2010). Involving business users in formal modeling using natural language pattern sentences. In *Proceedings of the 17th International Conference on Knowledge Engineering and Management by the Masses (EKAW'10)* (pp. 31-43). Berlin: Springer-Verlag. http://dx.doi.org/10.1007/978-3-642-16438-5_3.
- Walia, G. S., & Carver, J. C. (2009). A systematic literature review to identify and classify software requirement errors. *Information and Software Technology*, 51(7), 1087-1109. <http://dx.doi.org/10.1016/j.infsof.2009.01.004>.
- Yan, Z. (2007). Semantic business process modeling. In *Proceedings of the Knowledge Web PhD Symposium*. Innsbruck: KWEPSY.