

5 Regras Heurísticas

O conjunto de regras heurísticas apresentado neste capítulo apóia a interpretação dos resultados das medições na avaliação de sistemas. As regras (e o método apresentado no capítulo anterior) não assumem que a solução OA é sempre uma boa solução. Mesmo na presença de interesses transversais, as regras são abrangentes o suficiente para apontar problemas tanto em soluções OO quanto OA. Além disso, a abordagem também detecta, pelo uso de regras de acoplamento e coesão, os potenciais interesses transversais que não deveriam ser modularizados em aspectos. Cada regra é uma expressão baseada em métricas que indica certas características dos artefatos avaliados e podem apontar problemas potenciais. O principal objetivo das regras é fornecer informação mais abstrata ao desenvolvedor, evitando que este trabalhe diretamente com números e, portanto, facilitar a interpretação dos resultados. No contexto desta dissertação, o conjunto de regras pode ser usado na identificação de problemas nos artefatos de projeto detalhado e código. Tanto problemas de separação de interesses (SI), como de acoplamento e coesão podem ser apontados pelas regras heurísticas. Entretanto, o foco deste estudo é avaliação orientada a interesses e problemas de acoplamento e coesão só são relevantes quando estes problemas estão relacionados a uma ineficiente SI.

Como mencionado no capítulo anterior, a aplicação de regras heurísticas alerta o projetista sobre possíveis problemas causados por interesses transversais, no entanto, elas não garantem que os problemas existam realmente e que o software precise ser alterado. Portanto, mesmo quando as regras geram alertas, o projetista deve analisar o projeto e / ou código para verificar quais são as razões destes alertas. Essa tarefa representa a atividade de análise no método de avaliação apresentado no Capítulo 4. Na verdade, o benefício principal trazido pela utilização destas regras é que seus alertas contêm informações que ajudam o projetista a se concentrar em alguns artefatos do software que são possivelmente problemáticos. Depois da análise, se o projetista detectar que realmente existem

problemas causados por interesses transversais, ele pode reestruturar o sistema no sentido de melhor modularizar tais interesses. A reestruturação do software pode ser feito usando refatorações orientadas a aspecto como demonstrado na Subseção 4.2.4.

As regras heurísticas encontradas neste capítulo foram definidas a partir da experiência adquirida em estudos empíricos [8] [21] [27] que usam métricas. Estes estudos tanto avaliam a separação de interesses em sistemas quanto comparam projetos orientados a objetos (OO) a orientado a aspecto (OA) do mesmo sistema. Tais regras tornam-se muito importantes nas fases de projeto detalhado e código, devido ao grande número de elementos existentes nos artefatos (classes, métodos, atributos etc.) e, conseqüentemente, ao grande volume de dados a avaliar. Cada regra que pode ser expressa como um comando condicional usando a seguinte estrutura:

SE <condição> **ENTÃO** <conseqüência>

A *condição* envolve uma combinação de métricas e valores limites (Seção 5.4) para avaliação de interesses e artefatos, enquanto a *conseqüência* se apresenta como uma classificação do interesse ou artefato avaliado. O grupo de regras é dividido em duas categorias: regras de separação de interesses e regras de acoplamento e coesão. A primeira categoria de regras identifica problemas de projeto diretamente relacionados a uma imprópria modularização de interesses, enquanto a segunda categoria dá suporte à identificação de problemas de acoplamento e coesão que se relacionam aos problemas de SI. Apresentamos estas duas categorias de regras nas seções 5.2 e 5.3, respectivamente. Antes disso, na seção a seguir apontamos alguns problemas não triviais que podem ser resultantes de uma análise equivocada em atividades de medições. É importante ressaltar que estes problemas podem ser minimizados pela utilização do conjunto de regras heurísticas. Ao final deste capítulo são levantados alguns pontos relacionados à definição de valores limites para as regras.

5.1. Problemas de Interpretação em Medições

Uma grande dificuldade em se trabalhar com métricas é como lidar com seus resultados, ou seja, como toda aquela quantidade de dados pode ajudar na

melhoria da qualidade de software. Nesta seção examinamos alguns problemas que impactam a qualidade do sistema e estão relacionados a uma interpretação equivocada dos resultados de medições. Foram identificados seis destes problemas no decorrer do estudo: (a) alerta falso por interesse espalhado e entrelaçado; (b) alerta falso por alto acoplamento e/ou baixa coesão; (c) resultados não mostram um problema existente; (d) resultados não indicam onde está o problema; (e) conflitos entre valores medidos; e (f) métricas não relacionam os problemas existentes. Estes problemas, que ocorrem especialmente em avaliações orientadas a aspectos, são apresentados e exemplificados a seguir.

Problema A: Alerta falso por interesse espalhado e entrelaçado

O primeiro problema ocorre quando o resultado das métricas de SI adverte o desenvolvedor de um possível interesse que se espalha e entrelaça pelos artefatos do sistema. O desenvolvedor poderia pensar que se trata de um interesse transversal, entretanto, em uma análise mais cuidadosa verifica-se que, na verdade, tal interesse se encontra bem modularizado. Neste caso, os valores apontam um problema que não existe no sistema. Para ilustrar uma instância deste problema, é usado o padrão de projeto *Factory Method* [26].

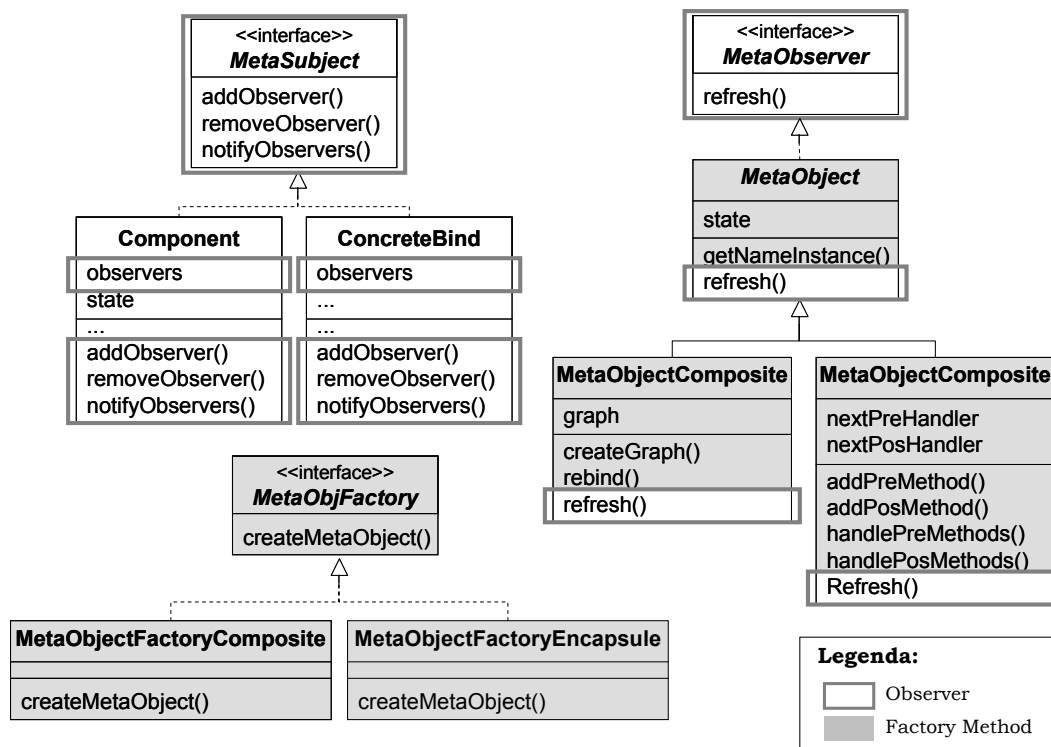


Figura 16 – Diagrama de classes OO destacando *Observer* e *Factory Method*

A Figura 16 apresenta o diagrama de classes parcial de um dos nossos estudos experimentais envolvendo um sistema de *middleware* reflexivo baseado no modelo *OpenOrb* [8] (Subseção 7.1.2). Neste diagrama são destacados os elementos que implementam os padrões *Factory Method* e *Observer* [26]. A aplicação de quatro métricas de SI neste subsistema resulta no conjunto de dados da Tabela 3. Estas métricas são CDC, CDLOC, NOAconcern e NOOconcern e encontram-se definidas na Tabela 1 da Subseção 4.2.1.

Tabela 3 – Resultado das métricas de SI para *Observer* e *Factory Method*

Componentes	Observer		Factory Method	
	CDC = 7	CDLOC = 18	CDC = 6	CDLOC = 8
	NOAconcern	NOOconcern	NOAconcern	NOOconcern
Component	1	3	0	0
ConcreteBind	1	3	0	0
MetaObject	0	1	1	1
MetaObjectComposite	0	1	1	2
MetaObjectEncapsu	0	1	2	4
MetaObjectFactoryComposite	0	0	0	1
MetaObjectFactoryEncapsule	0	0	0	1
MetaObjFactory	0	0	0	1
MetaObserver	0	1	0	0
MetaSubject	0	3	0	0

Em uma análise descuidada dos resultados apresentados na tabela acima, o desenvolvedor pode vir a acreditar que ambos os padrões são exemplos de interesses transversais – o que não é verdade para o padrão *Factory Method* [27] [36]. Esta interpretação errada é induzida pelos altos valores obtidos por estes padrões nas métricas. Em especial, os valores das métricas CDC e CDLOC para o padrão *Factory Method* são 6 e 8, respectivamente. Tais valores significam que seis componentes (dos 10 apresentados no diagrama) possuem código relativo ao padrão *Factory Method* e existem oito pontos em que há uma troca entre o interesse deste padrão e os outros interesses do sistema. Portanto, o padrão *Factory Method* encontra-se espalhado e entrelaçado pelos elementos do diagrama. Entretanto, isto não significa que este seja um interesse transversal, como na verdade não o é. Uma análise minuciosa do diagrama de classes (Figura 16) mostra que o padrão *Factory Method* encontra-se bem modularizado nos seis componentes em que ele se encontra. Além disso, o principal propósito destes seis componentes é implementar tal padrão [36]. Podemos verificar ainda, que existem alguns elementos do padrão *Observer* nos componentes que implementam o

padrão *Factory Method* causando entrelaçamento de interesses. Por todos esses fatores, pode-se concluir que os valores altos para o padrão *Factory Method* é resultado da incompleta modularização do padrão *Observer* e que as métricas não apresentam isto com clareza. Ou seja, caso o padrão *Observer* seja separado, o problema de entrelaçamento entre estes padrões é resolvido.

Problema B: Alerta falso por acoplamento e coesão

Na abordagem orientada a interesses proposta neste trabalho, um dos objetivos é conquistar uma melhoria em outros atributos do sistema, como acoplamento e coesão, a partir de uma melhor separação de interesses. Neste sentido, problemas de acoplamento e coesão somente são relevantes quando estes podem ser resolvidos por uma separação mais adequada dos interesses. Por outro lado, o elevado valor de algumas métricas de acoplamento e/ou coesão pode alertar o desenvolvedor de um possível problema. Em um caso menos grave, este alerta levaria o desenvolvedor a verificar se realmente existe algum problema que pode ser solucionado pela melhor SI. Em contrapartida, uma situação mais grave seria a busca de solução, possivelmente utilizando refatorações OA no projeto, quando o alerta não retrata problema grave ou que deva ser resolvido por técnicas de DSOA.

Tabela 4 – Resultado de acoplamento e coesão para o padrão *Factory Method*

Componentes	Factory Method			
	LCOO	CBC	DIT	NOC
ButtonCreator		3	2	0
ButtonCreator1		3	2	0
ButtonCreator2		3	2	0
GUIComponentCreator	3	6	1	6
LabelCreator		2	2	0
LabelCreator1		2	2	0
LabelCreator2		2	2	0
Main		1	1	0

Para ilustrar este problema de interpretação de números, utilizamos os resultados de medições feitas em uma implementação OO do padrão *Factory Method* [26] [36]. Esta implementação é uma extensão da versão proposta por Hannemann e Kiczales [36] e parte de nosso primeiro estudo experimental (Subseção 7.1.1). Os dados da Tabela 4 apresentam o resultado das medições neste estudo. Este resultado mostra valores elevados para o componente

`GUIComponentCreator` para as métricas LCOO, CBC e NOC em relação aos demais componentes do sistema. Em uma avaliação orientada a interesses, este fato pode levar o desenvolvedor a procurar uma forma de melhorar tais valores pela separação de interesses. Entretanto, o componente `GUIComponentCreator` não apresenta problema de SI, uma vez que seu único propósito é implementar o papel *Creator* do padrão *Factory Method*.

Problema C: Resultados não mostram o problema

Às vezes, as métricas não revelam os problemas existentes. Um caso típico ocorre quando um determinado interesse é conhecido e classificado como transversal, mas os resultados das métricas de SI não mostram problemas para este interesse. Este fato ocorre porque mesmo sendo um interesse transversal, pode haver situações em que não há um espalhamento e entrelaçamento muito intenso deste interesse. Uma instância do problema em que as métricas não mostram problemas aparece em nosso segundo estudo experimental (Subseção 7.1.2), quando avaliamos o padrão *Singleton* [26]. O diagrama de classes da Figura 17 destaca os elementos necessário para implementar este padrão e o padrão *Façade* [26] no estudo da *middleware* reflexiva baseada em *OpenOrb* [8] [9].

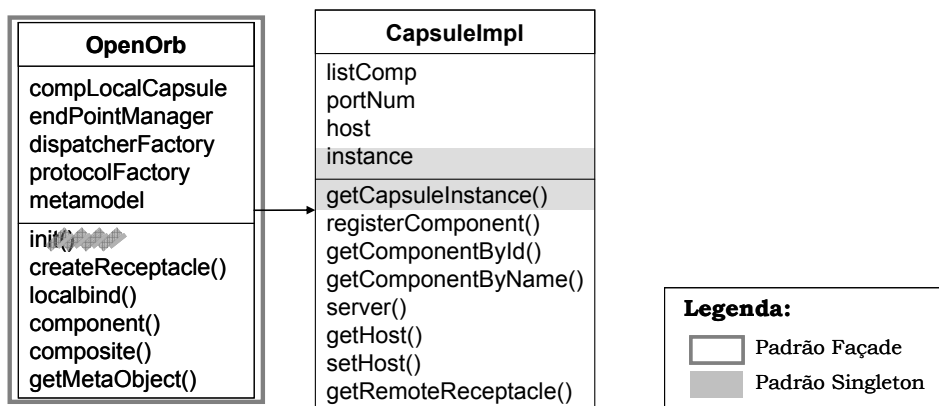


Figura 17 – Diagrama de classe OO destacando padrões *Façade* e *Singleton*

Tabela 5 – Resultado das métricas de SI para padrões *Façade* e *Singleton*

Componentes	Façade		Singleton	
	CDC = 1	CDLOC = 2	CDC = 2	CDLOC = 4
	NOAconcern	NOOconcern	NOAconcern	NOOconcern
CapsuleImpl	0	0	1	1
OpenOrb	5	6	0	0

O padrão *Singleton* [26] é classificado na literatura como um interesse transversal [9] [27] [28] [36], por dois motivos principais: (i) ele, tipicamente, se espalha pelos componentes e entrelaça aos outros interesses [27] e (ii) os componentes que implementam este interesse possuem outras funcionalidades na aplicação além de implementar o padrão [36]. Entretanto, olhando os dados da Tabela 5 temos a impressão de que o padrão *Singleton* se encontra bem modularizado, assim como o padrão *Façade*. Esta interpretação errada é resultado dos baixos valores apontados em todas as métricas de SI (CDC = 2, CDLOC=4, NOAconcern = 1, NOOconcern = 1). Note que mesmo o valor 4 da métrica CDLOC pode ser considerado baixo, pois este valor significa que existem apenas dois blocos de elementos que implementam tal interesse. Portanto, podemos concluir que neste caso as métricas de SI não mostram o problema de interesse transversal existente na aplicação.

Problema D: Onde está o problema?

Uma situação interessante relacionada à métricas de SI é que, se por um lado elas indicam um problema de interesse transversal, por outro, elas nem sempre mostram onde está o problema. Ou seja, os números não apontam quais partes do projeto ou implementação devem ser alteradas para resolver o problema. Um exemplo ocorre em nosso terceiro estudo experimental, Portalware [30], quando avaliado o interesse Colaboração. Este estudo é descrito na Subseção 7.1.3 e o resultado das medições é apresentado no Apêndice C. Em uma avaliação do resultado das métricas de SI, o interesse Colaboração é provavelmente classificado como um interesse transversal devido ao seu alto grau de espalhamento e entrelaçamento. Os valores apontados pelas métricas CDC = 15 e CDLOC = 14 confirmam esta intuição. É importante notar que o número de componentes do sistema é 60 (métrica VS no Apêndice C), ou seja, 25 % dos componentes possuem código relativo ao interesse avaliado.

A questão levantada é quantos e quais componentes devem ser alterados para separar este interesse e resolver o problema. Em uma impressão inicial, acreditaríamos ter que alterar os 15 componentes que possuem código do interesse Colaboração. Por outro lado, em uma análise mais aprofundada constatamos que 9 destes 15 componentes são totalmente dedicados à realização do interesse e que em apenas 6 existe entrelaçamento. Neste caso, o número de componentes que o

desenvolvedor precisa verificar a existência de problemas de SI reduz de 15 para 6. Apesar de todas estas informações serem extraídas dos resultados das medições, elas não são facilmente identificadas em uma análise *ad hoc* e só podem ser obtidas de um profundo raciocínio dos números de fina granularidade.

Problema E: Resultados conflitantes entre as métricas

A dura atividade de análise em resultados de medição pode se tornar ainda mais difícil quando os números apontam para caminhos contrários. Torna-se difícil classificar um interesse como transversal quando as métricas indicam espalhamento, mas não entrelaçamento, ou vice-versa. Em nosso estudo do Portalware (Subseção 7.1.3) este problema ocorre quando avaliamos o interesse Adaptação. Os dados deste estudo são detalhados no Apêndice C e podemos verificar uma contradição entre os valores das métricas $CDLOC = 14$ e $CDC = 3$. Enquanto a primeira métrica aponta um alto grau de entrelaçamento entre interesses, a segunda indica pouco espalhamento deste interesse. Mesmo olhando para outras métricas, como *NOAconcern* e *NOOconcern*, concluir se o interesse Adaptação pode ser considerado transversal neste sistema não é fácil.

Problema F: Métricas não relacionam os problemas

O último problema de medição levantado nesta seção é o fato de ser geralmente difícil relacionar problemas apontados pelas métricas de SI a problemas indicados em outras métricas. Este mapeamento é interessante neste estudo porque (i) estudos recentes mostram que outros atributos de software devem ser observados juntamente com separação de interesse na avaliação da qualidade [8] [21] [27] e (ii) o método de avaliação proposto nesta dissertação se baseia no princípio de que a melhoria de SI é acompanhada por ganhos em outros atributos, como acoplamento e coesão. É ilustrado a seguir como um interesse transversal pode levar à perda de coesão e aumento de acoplamento nos componentes afetados por este interesse.

A Tabela 6 apresenta o resultado de medições de SI para o papel *Subject* do padrão *Observer* [26] em uma implementação OO do padrão. Esta tabela também retrata valores de acoplamento e coesão medidos neste software que é parte do primeiro estudo experimental apresentado na Subseção 7.1.1. É interessante

observar que a maior parte dos componentes que implementam o papel *Subject* (*Point*, *Point1*, etc.) apresenta altos valores para as métricas de perda de coesão (LCOO) e de acoplamento (CBC). Uma análise detalhada da implementação mostra que estas classes apresentam maior perda de coesão porque precisam implementar métodos do padrão que não têm relação com os outros métodos da classe. Os métodos do padrão adicionam às classes funcionalidades para incluir, remover e notificar observadores. Além disso, o aumento na medição de acoplamento para as classes que desempenham o papel de *Subject* é justificado porque este papel é altamente acoplado aos componentes que desempenham papéis de *Observer* no mesmo padrão. Portanto, a conclusão é que o papel *Subject* leva a deterioração de outros atributos de qualidade do sistema, como coesão e acoplamento.

Tabela 6 – Resultado das métricas para o papel *Subject* do padrão *Observer*

Componentes	<i>Subject (Padrão Observer)</i>			
	CDC = 7	CDLOC = 56	VS = 12	
	NOAconcern	NOOconcern	LCOO	CBC
Observer	0	0		1
Point	1	3	15	4
Point1	1	3	15	4
Point2	1	3	15	4
Point3	1	3	15	4
Point4	1	3	15	4
Screen	1	3	1	3
Screen1	0	0	1	2
Screen2	0	0	1	2
Screen3	0	0	1	2
Screen4	0	0	1	2
Subject	0	3		1

5.2. Regras Heurísticas de Separação de Interesses

As regras apresentadas nesta seção utilizam uma combinação de métricas de separação de interesses (SI) e tamanho para avaliar os interesses do sistema e identificar possíveis problemas relacionados a interesses transversais. As métricas de SI utilizadas são CDC, CDLOC, NOAconcern, NOOconcern e LOCconcern; enquanto as métricas de tamanho são VS, NOA, NOO e LOC. Todas estas métricas são encontradas na Tabela 1 da Subseção 4.2.1. O objetivo principal do conjunto de regras de SI é classificar o interesse avaliado em “Modularizado”,

“Primário” ou “Secundário”. As duas primeiras classificações de interesses não indicam problemas de SI, enquanto a última pode ser problemática e, portanto, indicar que o projeto deva ser refatorado para melhor modularizar este tipo de interesse. Além das três categorias mencionadas, as regras também classificam os interesses em possivelmente primário, entrelaçado, de baixo espalhamento, de elevado espalhamento ou possivelmente secundário.

A Figura 18 retrata o diagrama com as possíveis classificações de um interesse e os relacionamentos de transição entre elas. As classificações são representadas por elipses, enquanto as transições são arcos direcionados ligando duas classificações. Além disso, as transições estão associadas às regras e estas definem quando ocorre uma mudança de categoria do interesse. As elipses do diagrama podem indicar três tipos de estados de classificação: inicial quando tracejado, intermediário quando representado em linha simples e final quando em linha dupla. Como discutimos no decorrer desta seção, um estado intermediário pode se tornar a classificação final do interesse quando nenhuma das transições que saem deste estado for validada pelas regras. O significado de cada uma das categorias presentes na Figura 18 é discutido a seguir.

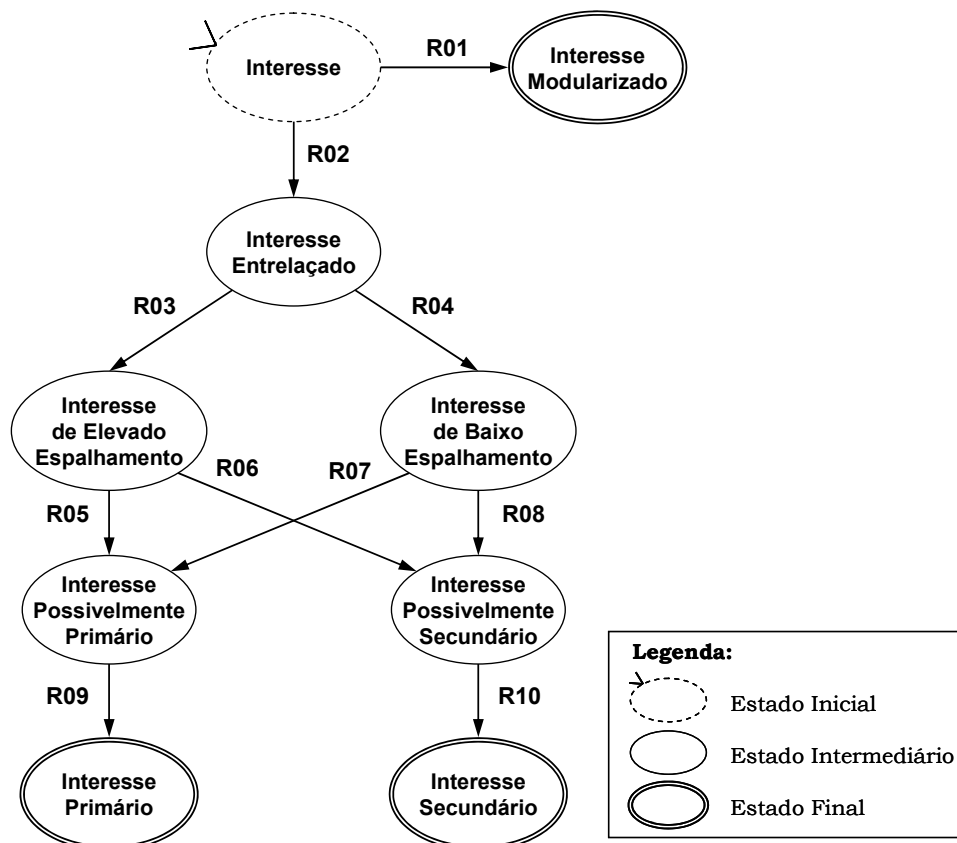


Figura 18 – Diagrama de estados dos interesses e regras de transição

Interesse Modularizado. Um interesse é classificado como *modularizado* quando todos os componentes responsáveis pela sua implementação são totalmente dedicados a este interesse. Ou seja, o interesse não possui entrelaçamento com nenhum outro interesse do sistema. Em um projeto ideal, todos os interesses deveriam estar totalmente separados e, portanto, seriam classificados como interesses modulares.

Interesse (Possivelmente) Primário. Em cada componente do sistema é definido um conjunto de dados e funções que, idealmente, atendem a um mesmo propósito. Este propósito é o que chamamos interesse primário do componente. Um interesse que é primário em um componente pode não ser primário em outros. Desta forma, um interesse somente é classificado como *primário* (no sistema) se ele é primário em todos os componentes em que ele se encontra. As regras classificam o interesse como *possivelmente primário*, quando este é um bom candidato a ser primário, mas não apresenta todas as características para tal.

Interesse Entrelaçado. Um interesse é classificado como *entrelaçado* quando este se encontra misturado a outros interesses dentro de um componente, ou seja, parte das operações e atributos do componente é destinada a um interesse e outra parte é dedicada a outro. O grau de entrelaçamento do interesse é fundamental para diferir sua classificação entre modularizado ou primário, pois, enquanto o “Interesse Modularizado” não se entrelaça a outros interesses em nenhum componente, o “Interesse Primário” pode se entrelaçar.

Interesse (Possivelmente) Secundário. Apesar de cada componente do sistema ser dedicado a um propósito principal, o “Interesse primário”, eles freqüentemente implementam outros interesses. Estes outros interesses são chamados secundários, pois incluem responsabilidades que não são as principais do componente. Mesmo o interesse que é primário em um componente pode ser secundário em outros. Assim, o interesse é classificado como *secundário* (no sistema) se ele é secundário em pelo menos um componente.

Interesse de Elevado Espalhamento e Interesse de Baixo Espalhamento. O espalhamento é a característica do interesse relacionada ao número de componentes afetados por ele. O grau de espalhamento de um interesse é variado e, portanto, são definidas duas classificações: (i) *elevado espalhamento*, quando vários componentes implementam o interesse; e (ii) *baixo espalhamento*, quando apenas alguns componentes são dedicados ao interesse. Quanto maior é o

espalhamento pelo sistema do interesse, maiores são as chances deste interesse ser transversal.

A Tabela 7 apresenta o conjunto de regras de separação de interesses que classifica os interesses de acordo com as categorias apresentadas. As regras que compõem este grupo devem ser aplicadas sequencialmente, na ordem em que elas são enumeradas, de tal forma que a classificação do interesse é gradativamente refinada. É importante ressaltar que as categorias não são exclusivas, ou seja, a nova classificação dada ao interesse por uma regra é cumulativa com a classificação dada pela regra anterior. Por exemplo, se a Regra 02 classifica um determinado interesse como “Entrelaçado”, a Regra 03 como “Elevado Espalhamento” e a Regra 06 como “Possivelmente Secundário”, então este interesse possui as três categorias. Em outras palavras, é um interesse “Entrelaçado, de Elevado Espalhamento e Possivelmente Secundário”. Como mencionado anteriormente, as regras utilizam as métricas de SI e tamanho listadas na Tabela 1 (Subseção 4.2.1).

Tabela 7 – Regras heurísticas de separação de interesses

R01	Se CDLOC é 2 então INTERESSE MODULARIZADO
R02	Se CDLOC é maior que 2 então INTERESSE ENTRELAÇADO
R03	Se CDC / VS de INTERESSE ENTRELAÇADO é alto então INTERESSE DE ELEVADO ESPALHAMENTO
R04	Se CDC / VS de INTERESSE ENTRELAÇADO não é alto então INTERESSE DE BAIXO ESPALHAMENTO
R05	Se (NOAconcern / NOA é alto) e (NOOconcern / NOO é alto) para todos os componentes de INTERESSE DE ELEVADO ESPALHAMENTO então INTERESSE POSSIVELMENTE PRIMÁRIO
R06	Se (NOAconcern / NOA é baixo) e (NOOconcern / NOO é baixo) para pelo menos um componente de INTERESSE DE ELEVADO ESPALHAMENTO então INTERESSE POSSIVELMENTE SECUNDÁRIO
R07	Se (NOAconcern / NOA é alto) e (NOOconcern / NOO é alto) para todos os componentes de INTERESSE DE BAIXO ESPALHAMENTO então INTERESSE POSSIVELMENTE PRIMÁRIO
R08	Se (NOAconcern / NOA é baixo) e (NOOconcern / NOO é baixo) para pelo menos um componente de INTERESSE DE BAIXO ESPALHAMENTO então INTERESSE POSSIVELMENTE SECUNDÁRIO
R09	Se (LOCconcern / LOC é alto) para todos os componentes de POSSÍVEL INTERESSE PRIMÁRIO então INTERESSE PRIMÁRIO
R10	Se (LOCconcern / LOC é baixo) para pelo menos um componente de POSSÍVEL INTERESSE SECUNDÁRIO então INTERESSE SECUNDÁRIO

As duas primeiras regras da Tabela 7 (R01 e R02) utilizam a métrica CDLOC para classificar o interesse em “Modularizado” ou “Entrelaçado”. Se o valor de CDLOC é dois, isto significa que o interesse está localizado em um único bloco sombreado. Mesmo que este interesse esteja em vários componentes, o valor 2 indica que estes componentes não possuem código relativo a nenhum outro interesse do sistema. Conseqüentemente, o interesse está modularizado e não existe entrelaçamento com outros interesses. Por outro lado, se CDLOC for maior que dois, o interesse avaliado se entrelaça a pelo menos um outro interesse. Isto é fácil de ser verificado, pois quando CDLOC é 4, existem dois blocos disjuntos sombreados para o interesse e, entre estes blocos, há um segundo interesse que causa o entrelaçamento.

As regras R03 e R04 são usadas para verificar se o interesse, quando classificado anteriormente como entrelaçado, se espalha por um grande número de componentes ou não. As métricas de CDC e VS provêm informações sobre este espalhamento através da medição do percentual de componentes afetados pelo interesse em relação ao número total de componentes no sistema (CDC/VS). Baseado nesta percentagem, o interesse é classificado em “Interesse de Baixo Espalhamento” ou “Interesse de Elevado Espalhamento”. O desenvolvedor deve estar consciente de interesses que se espalham por muitos componentes, pois qualquer alteração em algum destes interesses pode impactar em grande reestruturação do sistema.

Tanto um “Interesse de Elevado Espalhamento”, quanto um “Interesse de Baixo Espalhamento” pode ser primário ou secundário nos componentes do sistema. As regras R05 e R06 são usadas para decidir se um “Interesse de Elevado Espalhamento” é “Possivelmente Primário” ou “Possivelmente Secundário”, enquanto as regras R07 e R08 fazem análise semelhante para um “Interesse de Baixo Espalhamento”. O raciocínio aplicado nestas quatro regras é que um componente dedica a maior parte dos elementos internos para implementação de seu interesse primário. Desta forma, estas regras utilizam as métricas NOA, NOAconcern, NOO e NOOconcern para verificar a percentagem de atributos ($\text{NOAconcern} / \text{NOA}$) e de operações ($\text{NOOconcern} / \text{NOO}$) que cada componente dedica ao interesse avaliado. Se a percentagem de atributos e de operações é alta em todos os componentes, o interesse é classificado como “Possivelmente Primário”. Entretanto, se ambas as percentagens são baixas para

pelo menos um componente, o interesse é classificado como “Possivelmente Secundário”.

Completando o grupo de SI, as duas regras seguintes são responsáveis por classificar o interesse em “Primário” ou “Secundário”. O interesse “Possivelmente Primário” pode ser classificado como “Primário” pela regra R09, enquanto o interesse “Possivelmente Secundário” pode ser classificado como “Secundário” pela regra R10. Nestas duas últimas regras, o comportamento do interesse é inspecionado no contexto das linhas de código destinadas à sua implementação. O raciocínio é similar ao desenvolvido para as regras R05 a R08 e parte do princípio que um componente dedica a maior parte de suas linhas de código ao interesse primário. Portanto, a regra R09 classifica um interesse como “Primário” se este é “Possivelmente Primário” e a maior parte das linhas de código é destinada à implementação deste interesse nos componentes em que ele se encontra. Da mesma forma, um interesse é classificado como “Secundário” pela regra R10 se este é “Possivelmente Secundário” e, em pelo menos um componente, a maior parte das linhas de código não é destinada a sua implementação.

Mesmo o interesse classificado pela regra R09 como “Primário” se encontra entrelaçado a outros interesses (note que ele passou pela regra R02). Mas isto não significa que interesses primários sejam problemas de SI no sistema, pois, provavelmente o entrelaçamento é causado pelos outros interesses que são secundários. Um exemplo de interesse primário e entrelaçado pelo sistema é o padrão *Factory Method* [26] explorado na implementação da Seção 5.1 (Figura 16) em uma combinação com o padrão *Observer* [26]. Se naquele sistema todos os interesses forem sombreados e avaliados pelas regras, o padrão *Factory Method* seria classificado como “Interesse Primário” enquanto o *Observer* seria classificado como “Interesse Secundário”. Neste caso, o *Observer* é o interesse que possivelmente mereça ser melhor modularizado.

É importante notar que algumas das métricas utilizadas pelas regras da Tabela 7 só podem ser obtidas em artefatos de código e isto as torna restrita ao nível de implementação. As regras R09 e R10 são dependentes de métricas do nível de implementação e, portanto, estas regras não podem ser aplicadas caso os interesses estejam sendo avaliados no nível de projeto. Este fato justifica a divisão nos dois últimos níveis da avaliação. Se apenas artefatos de projeto detalhado estiverem disponíveis para a avaliação da separação de interesses do sistema, as

regras R01 a R08 podem ser utilizadas e os estados finais passam a ser “Interesse Possivelmente Primário” e “Interesse Possivelmente Secundário”.

5.3.

Regras Heurísticas de Acoplamento e Coesão

O principal objetivo das regras apresentadas na seção anterior é identificar “Interesses Secundários” e “Possivelmente Secundários”, pois estes podem representar oportunidades para refatorações OA. Além disso, estas regras de SI apontam os componentes que devem ser alterados para melhor separar tais interesses. Em outras palavras, a aplicação das regras de SI identifica um conjunto de componentes (possivelmente vazio) que possuem um ou mais “Interesses Secundários” ou “Possivelmente Secundários”. Desta forma, após a aplicação das regras de SI, um próximo passo na avaliação é identificar os possíveis efeitos destes interesses em outros atributos dos componentes do software, como acoplamento e coesão.

No decorrer dos estudos experimentais (Capítulo 7) foi verificado que a separação mais adequada dos interesses, em especial interesses secundários, tem como consequência a melhoria no acoplamento e coesão dos componentes do sistema. A melhora de coesão dos componentes após a retirada dos interesses secundários se justifica porque estes componentes possuíam responsabilidades não relacionadas. Ou seja, os componentes possuíam um conjunto de operações e atributos que implementava o interesse primário (ou propósito principal do componente) e outros conjuntos, provavelmente disjuntos, de operações e atributos que atendiam aos interesses secundários. Além disso, o acoplamento de um componente diminui após a retirada dos interesses secundários. Isto ocorre porque os componentes que implementam um mesmo interesse precisam se comunicar. Consequentemente, quando um componente é responsável por vários interesses, este também é responsável pelos relacionamentos com componentes destes interesses.

Nesta seção é apresentado um conjunto de regras de acoplamento e coesão que é complementar ao de SI (Seção 5.2). A aplicação deste segundo conjunto de regras constitui um passo seguinte às regras de SI e somente avaliam componentes que possuem algum “Interesse Secundário” ou “Possivelmente Secundário”. Ou seja, após a aplicação das regras de SI em todos os interesses do sistema, se alguns

deles forem classificados como “Secundário”, ou “Possivelmente Secundário”, as regras de acoplamento e coesão são utilizadas nos componentes em que exista tal interesse. Assim, o objetivo destas regras é auxiliar o desenvolvedor na identificação de componentes com problemas de acoplamento e coesão causados por interesses transversais. As regras utilizam as métricas CBC, LCOO e VS apresentadas na Tabela 1 (Subseção 4.2.1). A métrica VS aparece indiretamente nas regras, uma vez que ela é usada para calcular a média de acoplamento e perda de coesão dos componentes.

O conjunto de regras de acoplamento e coesão pode classificar os componentes em dois grupos em seu estado final: “Componente Candidato a Reestruturação Global” e “Componente Candidato a Extração de Interesses”. Além destas duas categorias, as regras de acoplamento e coesão também classificam os componentes nos seguintes estados intermediários: “Componente de Baixa Coesão”, “Componente de Elevada Coesão”, “Componente de Baixo Acoplamento” e “Componente de Elevado Acoplamento”. A Figura 19 apresenta o diagrama de categorias dos componentes e suas transições. Como no diagrama de interesses (Figura 18), neste diagrama as transições de estados também são definidas pelas as regras. A seguir são definidos os seis estados (finais e intermediários) da Figura 19.

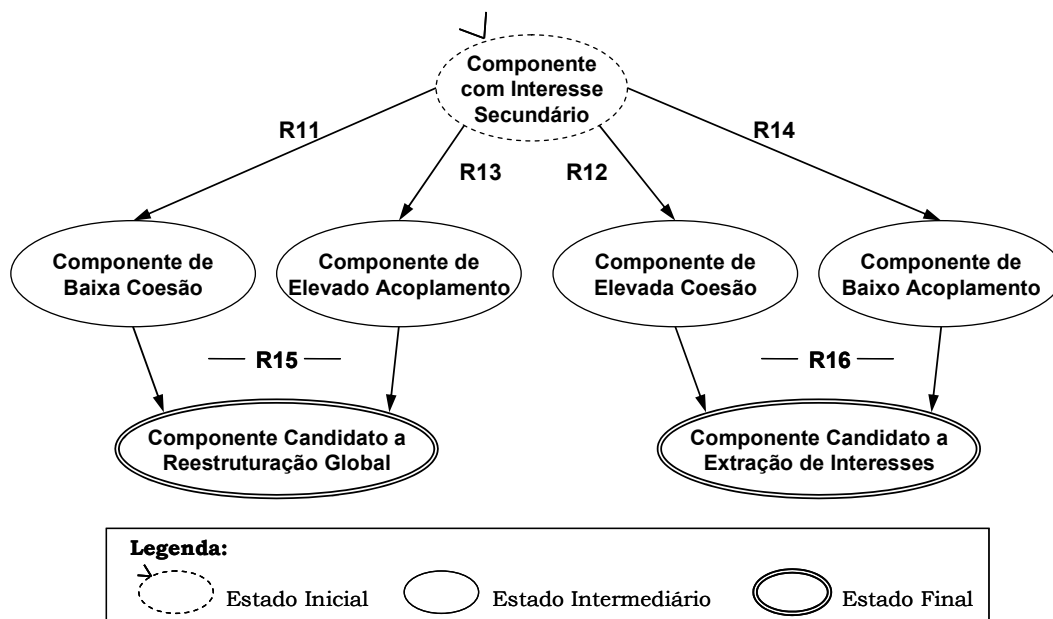


Figura 19 – Diagrama de estados dos componentes e regras de transição

Componente de Elevada Coesão e Componente de Baixa Coesão. A coesão interna de um componente é geralmente avaliada pelos relacionamentos existentes entre operações e atributos [14] [37]. Desta forma, as regras heurísticas propostas utilizam a métrica LCOO [14] [57] para classificar os componentes em *elevada coesão* ou *baixa coesão*. Um “Componente de Elevada Coesão” possui superior valor de coesão quando comparado à média apresentada pelos demais componentes do sistema. Em contrapartida, um “Componente de Baixa Coesão” apresenta valor inferior à média do sistema.

Componente de Elevado Acoplamento e Componente de Baixo Acoplamento. O acoplamento de um componente indica sua dependência em relação a outros componentes do sistema. Para avaliar o grau de acoplamento são verificadas as conexões com outros componentes, como chamadas a atributos e operações de outros componentes. No conjunto de regras proposto, o acoplamento é avaliado pela métrica CBC [14] [57] e os componentes podem ser classificados como *elevado acoplamento* ou *baixo acoplamento*. Um “Componente de Elevado Acoplamento” apresenta grande dependência de outros componentes, enquanto em um “Componente de Baixo Acoplamento” esta dependência é baixa. No contexto das regras apresentadas nesta seção, o acoplamento de um componente pode ser considerado baixo, ou alto, quando comparado com a média de acoplamento dos demais componentes do sistema.

Componente Candidato a Reestruturação Global. Quando um componente é classificado pelas regras como *candidato a reestruturação global* significa que este componente apresenta interesses secundários, baixa coesão e alto acoplamento. Estas características indicam problemas na definição do componente e, portanto, sugerem que este deva ser totalmente reestruturado. O principal objetivo das regras é encontrar oportunidades de aperfeiçoamento do sistema pela melhor separação de interesses, portanto, uma possível ação a ser tomada para resolver os problemas identificados pelas regras é utilizar refatorações OA para modularizar os interesses secundários destes componentes. É importante ressaltar que o “Componente Candidato a Reestruturação Global” representa o principal alerta de problema apontado pelas regras heurísticas deste trabalho.

Componente Candidato a Extração de Interesses. Um componente pode possuir interesses secundários e estes interesses não causarem problemas

aparentes de acoplamento e coesão. Quando isto ocorre, ou seja, quando as regras não detectam problemas nestes atributos dos componentes com interesses secundários, os componentes são classificados como *candidato a extração de interesses*. Esta classificação não descarta a existência de problemas nos módulos, apenas sugere que eles não são tão graves quanto na classificação anterior. Na verdade, o único possível problema verificado nestes componentes é a existência de interesse secundário e foi apontado pelas regras de SI. Quando comparado ao alerta feito pelas regras de SI “Interesse (Possivelmente) Secundário” ao alerta de “Componente Candidato a Extração de Interesses”, pode-se dizer que este último é menos grave, pois não foi detectado problema de acoplamento e coesão.

A Tabela 8 apresenta o conjunto de regras de acoplamento e coesão proposto neste trabalho. Estas regras são complementares às de SI (Seção 5.2) e classificam os componentes que possuem interesses secundários em uma das seis categorias mencionadas acima. Assim como ocorre nas regras de SI, a aplicação destas regras deve ser feita sequencialmente, na ordem em que elas são enumeradas. As regras de acoplamento e coesão utilizam as métricas CBC, LCOO e, indiretamente, VS para calcular as médias das duas métricas anteriores no sistema. O grupo é composto de 6 regras, numeradas de R11 a R16, que discutimos a seguir.

Tabela 8 – Regras heurísticas de acoplamento e coesão

R11	Se LCOO do COMPONENTE COM INTERESSE SECUNDÁRIO é alto em relação à média de LCOO dos componentes do sistema então COMPONENTE DE BAIXA COESÃO
R12	Se LCOO do COMPONENTE COM INTERESSE SECUNDÁRIO é baixo em relação à média de LCOO dos componentes do sistema então COMPONENTE DE ELEVADA COESÃO
R13	Se CBC do COMPONENTE COM INTERESSE SECUNDÁRIO é alto em relação à média de CBC dos componentes do sistema então COMPONENTE DE ELEVADO ACOPLAMENTO
R14	Se CBC do COMPONENTE COM INTERESSE SECUNDÁRIO é baixo em relação à média de CBC dos componentes do sistema então COMPONENTE DE BAIXO ACOPLAMENTO
R15	Se (COMPONENTE DE BAIXA COESÃO) e (COMPONENTE DE ELEVADO ACOPLAMENTO) então COMPONENTE CANDIDATO A REESTRUTURAÇÃO GLOBAL
R16	Se (COMPONENTE DE ELEVADA COESÃO) e (COMPONENTE DE BAIXO ACOPLAMENTO) então COMPONENTE CANDIDATO A EXTRAÇÃO DE INTERESSES

As duas primeiras regras da Tabela 8, R11 e R12, avaliam a coesão dos componentes que possuem interesses secundários. Estas regras utilizam as

métricas LCOO e VS para classificar os componentes em “Baixa Coesão” ou “Elevada Coesão”. A métrica VS aparece implicitamente na regra para o cálculo de média de perda de coesão do sistema, ou seja, o somatório da perda de coesão de todos os componentes dividido por VS. Um componente é classificado como “Componente de Baixa Coesão” quando seu valor de LCOO é maior que a média do sistema, enquanto um “Componente de Elevada Coesão” possui valor abaixo da média. Note que a métrica LCOO mede a perda de coesão, portanto, quanto maior o valor desta métrica, menor é a coesão do componente.

As regras R13 e R14 fazem uma avaliação semelhante às regras anteriores, porém, considerando o critério de acoplamento. Os componentes que possuem algum interesse secundário podem ser classificados em “Baixo Acoplamento” ou “Elevado Acoplamento” por estas regras. Na primeira categoria, estão os componentes que possuem valor de CBC abaixo da média do sistema e na segunda categoria os componentes com valor acima da média. Tanto as regras R11 e R12 quanto as regras R13 e R14 criam dois grupos complementares de componentes. A diferença está no critério utilizado para alocar os componentes nos grupos, coesão para o primeiro par de regras e acoplamento para o segundo par.

As últimas duas regras da Tabela 8, R15 e R16, não utilizam nenhum resultado de medição diretamente. O principal propósito destas regras é unificar estados nos quais os componentes apresentam problemas de acoplamento e coesão (R15) e estados nos quais os componentes não apresentam problemas nestes atributos (R16). Quando a aplicação das regras anteriores detecta que um componente é pouco coeso (R11) e muito acoplado (R13), este componente é classificado como “Candidato a Reestruturação Global” pela regra R15. Em contrapartida, o componente muito coeso (R12) e pouco acoplado (R14) é classificado como “Candidato a Extração de Interesses” pela regra R16. É importante ressaltar que em ambos os casos, os componentes possuem interesses secundários identificados durante a aplicação das regras de SI.

5.4. Definição de Valores Limite

Nesta seção discutimos um ponto importante em relação às regras heurísticas apresentadas no decorrer do capítulo, os valores limites. Tanto regras

de separação de interesses quanto regras de acoplamento e coesão dependem da definição de valores limites que, quando ultrapassados, disparam as conseqüências destas regras. Os valores limite são parâmetros importantes para a eficácia das regras e, portanto, devem ser muito bem definidos. Quando a conseqüência da regra depende de um valor abaixo do limite e este limite possui um valor acima do ideal, o resultado é muitos falsos positivos. Por outro lado, se o limite configurado na regra está abaixo do ideal, o resultado são muitos falsos negativos. O problema de encontrar o valor limite ideal não é novo e intrinsecamente caracteriza qualquer abordagem baseada em medição [40] [47].

Duas estratégias podem ser usadas na definição de valores limites: boas sugestões encontradas na literatura e configuração empírica. No primeiro caso, a escolha do valor limite é guiada por experiências semelhantes do passado e acertos apontados por autores de trabalhos relacionados a medições [47]. Quando estes acertos não estão disponíveis, os valores podem ser configurados de acordo com as características do sistema (e.g. tamanho) e avaliados em estudos empíricos. Por exemplo, se um determinado valor que foi definido gerar muitos falsos positivos ou falsos negativos do que resultados corretos, este valor pode ser alterado e re-avaliado. Na verdade, quando uma ferramenta de suporte às regras se encontra disponível, o ajuste do valor limite para os requisitos específicos de um projeto ou de um desenvolvedor é facilitado. Com uma ferramenta o desenvolvedor pode definir um valor e verificar imediatamente os efeitos desta ação (alertas excessivos ou escassos); variando continuamente este valor até restringir os alertas a um número que pode ser verificado em tempo hábil e verificar se os alertas gerados estão sendo efetivos.

Tabela 9 – Valores limite utilizados nas regras heurísticas

Regras	Valores Limites
Regra 01 e 02	2
Regras 03 e 04	50 % (se $VS \leq 10$) 40 % (se $10 < VS \leq 20$) 30 % (se $20 < VS \leq 50$) 20 % (se $50 < VS \leq 80$) 15 % (se $80 < VS$)
Regras 05, 06, 07 e 08	50 % (de NOA) 50 % (de NOO)
Regras 09 e 10	50 % (de LOC)
Regras 11 e 12	20 % (da média de LCOO)
Regras 13 e 14	20 % (da média de CBC)
Regras 15 e 16	-

Na abordagem desta dissertação são utilizadas ambas as estratégias para definição de valores limites, sugestões encontradas na literatura e configuração empírica. O ponto de partida são acertos de trabalhos anteriores e, então, os valores são customizados empiricamente para as características individuais de cada projeto. A Tabela 9 apresenta os valores limites para as dezesseis regras propostas neste capítulo e que foram utilizados nos estudos de caso realizados. As regras 15 e 16 são as únicas que não necessitam de valores limites, pois estas regras apenas unificam dois estados de classificação dos componentes (seção 5.3).

Como pode ser notado pela Tabela 9, o valor limite das regras 3 e 4 varia entre 15 % e 50 % dependendo do tamanho do projeto. Ou seja, o valor limite é de 15 % se o sistema avaliado possui mais de 80 componentes, 50 % se possui menos de 10 componentes e outro valor intermediário quando o número de componentes é maior que 10 e menor que 80. Esta variação de valor limite é feita para tentar acompanhar a intuição humana em relação ao alto espalhamento ou baixo espalhamento de um interesse. Por exemplo, em um sistema grande de 100 componentes quando um interesse atinge mais de 15 componentes deste sistema (portanto, mais de 15 %), este interesse será provavelmente classificado como de elevado espalhamento por um usuário humano (e também pelas regras). Por outro lado, em um sistema de 10 componentes é necessário que pelo menos 5 deles (50 %) possuam um determinado interesse para que o interesse seja considerado de elevado espalhamento pela maioria dos programadores.