

Universidade de São Paulo - USP
Universidade Federal de São Carlos - UFSCar
Universidade Estadual Paulista - UNESP

Heurísticas de Poda de Sentenças para a Sumarização Automática de Textos UNL: Estudo de Casos

Camilla Brandel Martins
Lucia Helena Machado Rino

NILC-TR-02-11

Julho, 2002

Série de Relatórios do Núcleo Interinstitucional de Linguística Computacional
NILC - ICMC-USP, Caixa Postal 668, 13560-970 São Carlos, SP, Brasil

Resumo

Este relatório apresenta um estudo de casos conduzido para testar o conjunto de heurísticas para sumarização de textos UNL implementadas no protótipo UNLSumm. Este protótipo recebe como entrada textos convertidos para a *Universal Networking Language*, UNL e, por meio das heurísticas de sumarização, gera um sumário também em UNL. Como estudos preliminares mostraram que alguns dos sumários gerados eram de má qualidade, levantamos algumas hipóteses sobre as heurísticas e conduzimos o estudo de casos para testá-las, buscando assim melhorar os resultados. Com base nesse estudo, aprimoramos o conjunto de heurísticas e o método utilizado para aplicá-las.

Este trabalho conta com o apoio financeiro da CAPES.



Índice

1	INTRODUÇÃO.....	1
2	HEURÍSTICAS PARA SUMARIZAÇÃO DE TEXTOS UNL	2
2.1	A INTERLÍNGUA UNL	2
2.2	O UNLSUMM NO CENÁRIO DO SISTEMA UNL DE COMUNICAÇÃO MULTILINGUAL	3
2.3	UNLSUMM: O MECANISMO DE PODA DAS HEURÍSTICAS DE SUMARIZAÇÃO	5
3	O ESTUDO DE CASOS	6
3.1	PRINCIPAIS MÉTRICAS PARA AVALIAÇÃO	6
3.1.1	<i>Textualidade</i>	6
3.1.2	<i>Preservação do Foco</i>	6
3.1.3	<i>Score dos Sumários</i>	7
3.1.4	<i>Taxa de Compressão</i>	7
3.2	PREMISSAS E HIPÓTESES DO ESTUDO DE CASOS	7
3.2.1	<i>Premissas</i>	8
3.2.2	<i>Hipóteses a Serem Verificadas</i>	8
3.3	METODOLOGIA DE ELABORAÇÃO DO ESTUDO DE CASOS	9
3.3.1	<i>Casos 1 e 2</i>	10
3.3.2	<i>Casos 3 e 4</i>	10
3.3.3	<i>Casos 5 e 6</i>	10
4	RESULTADOS OBTIDOS COM O ESTUDO DE CASOS	12
4.1	ESTUDO DE CASOS PARA O CORPUS UNU	12
4.1.1	<i>Aplicação individual de heurísticas</i>	12
4.1.1.1	Caso 1 – Heurísticas do Grupo A	12
4.1.1.2	Caso 2 – Heurísticas do Grupo B	13
4.1.1.3	Crítica Parcial	15
4.1.2	<i>Aplicação conjunta de heurísticas</i>	16
4.1.2.1	Caso 3 – Heurísticas do Grupo A	16
4.1.2.2	Caso 4 – Heurísticas do Grupo B	17
4.1.2.3	Crítica Parcial da Aplicação Conjunta de Heurísticas dos Grupos A e B	19
4.1.3	<i>Aplicação de Heurísticas Segundo sua Ordem</i>	19
4.1.3.1	Caso 5 – Aplicação de Heurísticas do Grupo A, depois do Grupo B	19
4.1.3.2	Caso 6 – Aplicação de Heurísticas do Grupo B, depois do Grupo A	21
4.1.3.3	Crítica Parcial Sobre a Ordem de Aplicação de Heurísticas dos Grupos A e B	22
4.1.4	<i>Crítica Geral Sobre o Corpus UNU</i>	22
4.2	ESTUDO DE CASOS PARA O CORPUS THESES	25
4.2.1	<i>Aplicação individual de heurísticas</i>	25
4.2.1.1	Caso 1 – Heurísticas do Grupo A	25
4.2.1.2	Caso 2 – Heurísticas do Grupo B	27
4.2.1.3	Crítica Parcial	29
4.2.2	<i>Aplicação conjunta de heurísticas</i>	29
4.2.2.1	Caso 3 – Heurísticas do Grupo A	29
4.2.2.2	Caso 4 – Heurísticas do Grupo B	30
4.2.2.3	Crítica Parcial da Aplicação Conjunta de Heurísticas dos Grupos A e B	32
4.2.3	<i>Ordem de Aplicação</i>	32
4.2.3.1	Caso 5 – Aplicar primeiro heurísticas do Grupo A	32
4.2.3.2	Caso 6 – Aplicar primeiro heurísticas do Grupo B	33
4.2.3.3	Crítica Parcial sobre a Ordem de Aplicação de Heurísticas dos Grupos A e B	34
4.2.4	<i>Crítica Geral Sobre o Corpus Theses</i>	35
5	CONCLUSÕES GERAIS DO ESTUDO DE CASOS.....	37
6	READEQUAÇÃO DO CONJUNTO DE HEURÍSTICAS.....	41
7	DISCUSSÃO FINAL	44

REFERÊNCIAS BIBLIOGRÁFICAS.....	46
ANEXO 1 – VERSÃO 1.3 DAS HEURÍSTICAS DE SUMARIZAÇÃO.....	47
ANEXO 2 – VERSÃO 2.0 DAS HEURÍSTICAS DE SUMARIZAÇÃO.....	50

Índice de Tabelas

TABELA 1: DESCRIÇÃO DOS PRINCIPAIS RLS NO UNLSUMM	3
TABELA 2: MEDIDAS DE AVALIAÇÃO DE SUMÁRIOS UNL	7
TABELA 3: ESTATÍSTICA DE PODA PARA O CASO 1	12
TABELA 4: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 1	13
TABELA 5: ESTATÍSTICAS DE PODA PARA O CASO 2	14
TABELA 6: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 2	14
TABELA 7: ESTATÍSTICAS DE PODA PARCIAIS PARA O CASO 3	16
TABELA 8: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 3	16
TABELA 9: ESTATÍSTICAS DE PODA PARCIAIS PARA O CASO 4	18
TABELA 10: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 4	18
TABELA 11: ESTATÍSTICAS DE PODA PARCIAIS PARA O CASO 5	20
TABELA 12: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 5	20
TABELA 13: ESTATÍSTICAS DE PODA PARCIAIS PARA O CASO 6	21
TABELA 14: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 6	21
TABELA 15: ESTATÍSTICA DE PODA GLOBAL PARA O CORPUS UNU	24
TABELA 16: ESTATÍSTICAS DE PODA PARA O CASO 1	26
TABELA 17: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 1	26
TABELA 18: ESTATÍSTICAS DE PODA PARA O CASO 2	28
TABELA 19: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 2	28
TABELA 20: ESTATÍSTICAS DE PODA PARCIAIS PARA O CASO 3	29
TABELA 21: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 3	29
TABELA 22: ESTATÍSTICAS DE PODA PARCIAIS PARA O CASO 4	31
TABELA 23: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 4	31
TABELA 24: ESTATÍSTICAS DE PODA PARCIAIS PARA O CASO 5	32
TABELA 25: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 5	32
TABELA 26: ESTATÍSTICAS DE PODA PARCIAIS PARA O CASO 6	33
TABELA 27: RESULTADOS PARA PRESERVAÇÃO DO FOCO E TEXTUALIDADE NO CASO 6	34
TABELA 28: ESTATÍSTICAS DE PODA GLOBAL PARA O CORPUS THESES	35
TABELA 29: ESTATÍSTICAS PARA SENTENÇAS COM TAXAS DE COMPRESSÃO ACIMA DE 70%	36
TABELA 15: ESTATÍSTICA DE PODA GLOBAL PARA O CORPUS UNU	38
TABELA 28: ESTATÍSTICAS DE PODA GLOBAL PARA O CORPUS THESES	38

Índice de Figuras

FIGURA 1: SISTEMA DE DECODIFICAÇÃO UNL-PORTUGUÊS	4
FIGURA 2: UNLSUMM ACOPLADO AO SISTEMA DE DECODIFICAÇÃO UNL-PORTUGUÊS	5

Índice de Gráficos

GRÁFICO 1: DESEMPENHO DOS SUMÁRIOS DO CORPUS UNU EM RELAÇÃO À TAXA DE COMPRESSÃO	25
GRÁFICO 2: DESEMPENHO DOS SUMÁRIOS EM RELAÇÃO À TAXA DE COMPRESSÃO	36
GRÁFICO 3: TAXA DE COMPRESSÃO EM RELAÇÃO AO SCORE	39
GRÁFICO 4: PRESERVAÇÃO DO FOCO E TEXTUALIDADE EM SUMÁRIOS PRODUZIDOS NOS CASOS 2,4 E 6	39
GRÁFICO 5: COMPARAÇÃO GERAL DOS SCORES OBTIDOS PARA O CORPUS UNU	41
GRÁFICO 6: COMPARAÇÃO GERAL DOS SCORES OBTIDOS PARA O CORPUS THESES	41
GRÁFICO 7: PERCENTUAL DE SUMÁRIOS DO CORPUS UNU QUE PRESERVAM O FOCO	42
GRÁFICO 8: PERCENTUAL DE SUMÁRIOS DO CORPUS UNU QUE PRESERVAM A TEXTUALIDADE	42
GRÁFICO 9: PERCENTUAL DE SUMÁRIOS DO CORPUS THESES QUE PRESERVAM O FOCO	42
GRÁFICO 10: PERCENTUAL DE SUMÁRIOS DO CORPUS THESES QUE PRESERVAM A TEXTUALIDADE	43

1 Introdução

O trabalho descrito neste relatório é parte de um projeto que visa a sumarização automática de textos UNL (Martins, 2001), i.e., textos codificados na *Universal Networking Language* (Uchida, 2000). Nesse projeto, a sumarização é realizada por meio de heurísticas de poda de textos UNL, descritas detalhadamente em (Martins e Rino, 2002). Um protótipo, denominado *UNL Summarizer* ou, simplesmente, UNLSumm, foi desenvolvido para automaticamente aplicar essas heurísticas a textos UNL, gerando assim, sumários UNL. No entanto, em estudos preliminares com o protótipo, verificamos que a qualidade de alguns sumários gerados era insatisfatória, pois eles não preservavam a idéia do texto original, tampouco permitiam a obtenção de sentenças gramaticais.

Assim, estabelecemos algumas hipóteses sobre os motivos que levaram a esses resultados e conduzimos o estudo de casos apresentados neste relatório para testá-las. De um modo geral, o estudo consistiu na geração de sumários utilizando o protótipo UNLSumm e, em seguida, na avaliação dos sumários gerados. Com base nos resultados obtidos, o conjunto de heurísticas foi remodelado de forma a melhorar o desempenho do UNLSumm.

No próximo capítulo descrevemos brevemente as características principais das heurísticas de sumarização e seu cenário de aplicação. Em seguida descrevemos o estudo de casos, incluindo as suas hipóteses e as métricas utilizadas (Capítulo 3). No Capítulo 4 apresentamos os resultados obtidos ao considerar alguns corpora de teste, apresentando as conclusões gerais no Capítulo 5. No capítulo 6 discutimos a readequação do conjunto de heurísticas e, finalmente, no Capítulo 7, apresentamos algumas considerações finais sobre o desempenho do UNLSumm .

2 Heurísticas para Sumarização de Textos UNL

As heurísticas de sumarização de textos UNL têm por objetivo condensá-los, produzindo textos UNL mais curtos, aqui chamados de sumários UNL. Elas se baseiam em características da interlíngua UNL, que define que um texto pode ser representado como um conjunto de suas sentenças, sendo que cada sentença de um texto UNL é, por sua vez, representada por um conjunto de relações semânticas, expressas por relações binárias, aqui referidas simplesmente como RBs. Assim, para sumarizar um texto completo, consideraremos a sumarização de cada uma de suas sentenças, pela aplicação de uma ou mais heurísticas, as quais tentam identificar as informações que possam ser consideradas não relevantes no contexto em foco, pela identificação das relações binárias correspondentes a essas informações. Tais RBs serão, em seguida, excluídas da sentença UNL em que estão inseridas, gerando assim o *sumário UNL*. Caracterizadas, portanto, pela exclusão de RBs, as heurísticas são chamadas de *heurísticas de poda*.

Como as heurísticas são baseadas na UNL, descrevemos, a seguir, as principais características dessa interlíngua, em especial, aquelas mais relevantes para a sumarização de textos UNL. Descrevemos, também, o modo como o UNLSumm pode ser utilizado no sistema de comunicação interlingual do Projeto UNL para, então, apresentar o mecanismo de poda, propriamente dito. Uma descrição completa do conjunto de heurísticas pode ser encontrada em (Martins e Rino, 2002).

2.1 A Interlíngua UNL

Tendo sido desenvolvida no Projeto UNL¹ para fins de comunicação interlingual, a UNL permite representar o significado de textos sentença por sentença, pretensamente de forma não ambígua. Cada sentença, por sua vez, é representada por um conjunto de relações binárias, ou simplesmente RBs, na forma $RL(UW_1, UW_2)$, sendo *RL*, ou *Relation Label* um rótulo que indica a relação semântica entre os dois conceitos universais UW_1 e UW_2 , aqui representando *Universal Words*.

Os RLS são expressos por mnemônicos de três letras. Por exemplo, o RL ‘plc’ pode relacionar uma ação ao local onde ela ocorre, enquanto ‘agt’ relaciona uma ação ao seu agente. A UNL é composta atualmente por 41 RLS distintos, no entanto, nem todos são utilizados em nosso modelo de sumarização, pois alguns claramente indicam informações essenciais. Por exemplo, enquanto uma informação de localização (expressa por ‘plc’ na UNL) pode ser considerada pouco relevante, dificilmente a informação sobre o agente de uma ação (expressa por ‘agt’ na UNL) será pouco importante em um evento. A Tabela 1 descreve os RLS que foram usados no estudo de casos, para evidenciar, em contextos específicos, a razão de alguma RB ser considerada supérflua e, portanto, passível de exclusão da sentença correspondente. As definições foram generalizadas com base em suas principais interpretações, a partir de análise de corpora de textos. Entretanto, a especificação original da UNL é bem mais abrangente, sendo que sua descrição completa pode ser encontrada em (Uchida, 2000).

¹ Uma descrição completa deste projeto e seus módulos pode ser encontrada em <http://www.unl.ias.unu.edu>.

<i>RL</i>	<i>Nome referência</i>	<i>Descrição</i>
agt	Agente	Descreve o agente que inicia uma ação
aoj	Atributo de objeto	Descreve o atributo de um elemento ²
ben	Beneficiário	Descreve o beneficiário ou vítima de uma ação, evento ou estado.
con	Condição	Descreve um evento ou estado que condiciona o evento em foco.
dur	Duração	Descreve o período de tempo durante o qual um evento ocorre ou um estado existe.
lpl	Lugar lógico	Descreve o cenário, ou lugar lógico onde uma ação ou evento ocorre.
man	Modo	Descreve um modo específico de realização de uma ação ou uma característica de um evento.
met	Método	Descreve o método necessário para a realização de uma ação ou evento.
mod	Modificador	Descreve um modificador de um conceito universal, restringindo a interpretação desse elemento.
nam	Nome	Descreve o nome próprio de um elemento ou objeto.
obj	Objeto	Descreve o elemento conceituado como diretamente afetado por uma ação ou evento.
plc	Local	Descreve o lugar onde uma ação ocorre, podendo ser físico ou lógico.
pos	Possuidor	Descreve o elemento que possui um objeto ou lugar.
ppl	Lugar físico	Descreve o lugar físico onde uma ação ocorre.
pur	Propósito	Descreve o propósito ou objetivo de uma ação ou evento; ou objetivo de algo existir.
qua	Quantidade	Descreve uma quantidade em relação a uma unidade de medida.
rsn	Razão	Descreve a razão de um evento acontecer ou um estado existir.
scn	Cenário	Descreve um mundo ou lugar virtual onde uma ação ou evento ocorre.
src	Fonte	Descreve o estado inicial de um elemento antes de um evento ou estado.
tim	Tempo	Indica uma informação temporal sobre uma ação ou evento.

Tabela 1: Descrição dos principais RLs no UNLSumm

2.2 O UNLSumm no cenário do Sistema UNL de comunicação multilingual

Como o UNLSumm consiste em um módulo de poda de códigos UNL, consideramos que ele pode ser inserido na própria arquitetura do Sistema UNL, conforme descrevemos nesta seção.

O Sistema UNL é caracterizado por permitir a comunicação multilingual fazendo uso da representação interlíngua, i.e., da própria UNL. Assim, é possível que, a partir de qualquer texto T1 em uma língua natural LN1, seja possível gerar seu correspondente texto T2 em uma língua natural LN2. Para tanto, são usadas duas ferramentas básicas: 1) a de codificação LN1-UNL, resultando na representação interlingual, ou código UNL, de T1 e 2) a de decodificação UNL-LN2, resultando em T2. No Sistema UNL, à primeira

² Utilizaremos o termo ‘elemento’ como tendo o significado de pessoa, objeto ou qualquer ser ou coisa que, em uma sentença, apareça como volitiva, ou seja, como tendo vontades e iniciativa de ação.

ferramenta deu-se o nome de EnCo; à segunda, o nome de DeCo (Uchida, 1997). Como o UNLSumm contempla diretamente códigos UNL, consideramos que estes são o resultado da codificação de um língua natural qualquer em UNL, ou seja, consideramos que a codificação LN-UNL foi previamente realizada, resultando no código UNL de nosso interesse. Assim, o módulo de poda será acoplado, diretamente, ao ambiente de decodificação.

Notamos que o DeCo é disponibilizado no Projeto UNL por seus mentores em sua versão genérica, devendo ser customizado para cada língua natural por meio de recursos lingüísticos. A Figura 1 ilustra o ambiente de decodificação quando o DeCo é customizado para o português (neste caso, referimo-nos à decodificação UNL - Português), aqui representado por PB (Português do Brasil), sendo de responsabilidade do NILC³ desenvolver os recursos lingüísticos que o customizam para o português. Estes incluem, basicamente, uma gramática (UNL-PB) e um dicionário bilíngüe (UNL-PB), conforme ilustra a Figura 1.

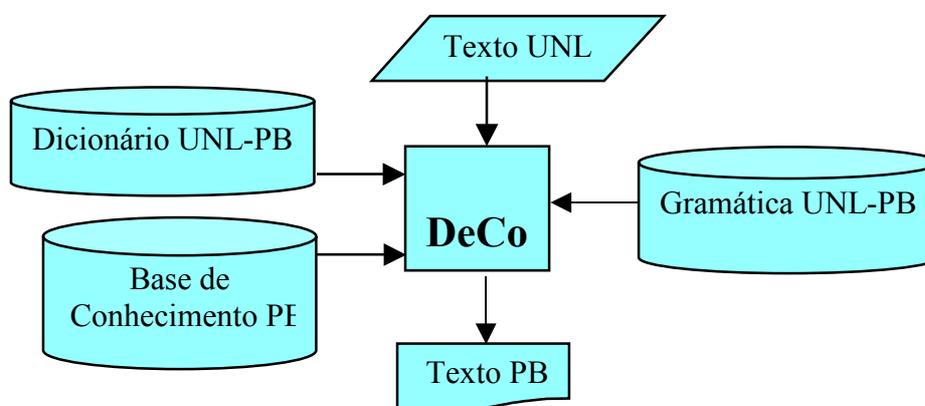


Figura 1: Sistema de Decodificação UNL-Português

O aproveitamento do cenário de decodificação UNL-Português ilustrado acima, para a sumarização automática de textos UNL, consiste em acoplar o UNLSumm, nosso mecanismo de poda de textos UNL, de forma que seus resultados – sumários UNL – possam ser decodificados. Esse acoplamento é ilustrado na Figura 2, resultando em um novo ambiente que pode ser explorado automaticamente de duas formas: ou gerando-se um texto em português a partir de um código completo (neste caso, estaremos supondo que houve a codificação LN1-UNL para, então, termos a decodificação UNL-PB) ou gerando-se um sumário em português, do suposto texto na LN1 (e, neste caso, o texto UNL original seria primeiro condensado, para que a decodificação se dê, agora, a partir do sumário UNL).

A função do UNLSumm, nesse novo ambiente, será, assim, aplicar o mecanismo de poda em um texto UNL considerado completo, para gerar um texto, ou sumário UNL, que será decodificado usando os mesmos recursos do português, já definidos no sistema de decodificação. Vale lembrar que, por contemplar somente códigos UNL, o UNLSumm pode, em princípio, ser acoplado a qualquer sistema de decodificação, ou seja, pode ser um sistema de sumarização para gerar textos em qualquer língua-destino LN2.

³ Núcleo Interinstitucional de Lingüística Computacional (<http://nilc.icmc.sc.usp.br>).

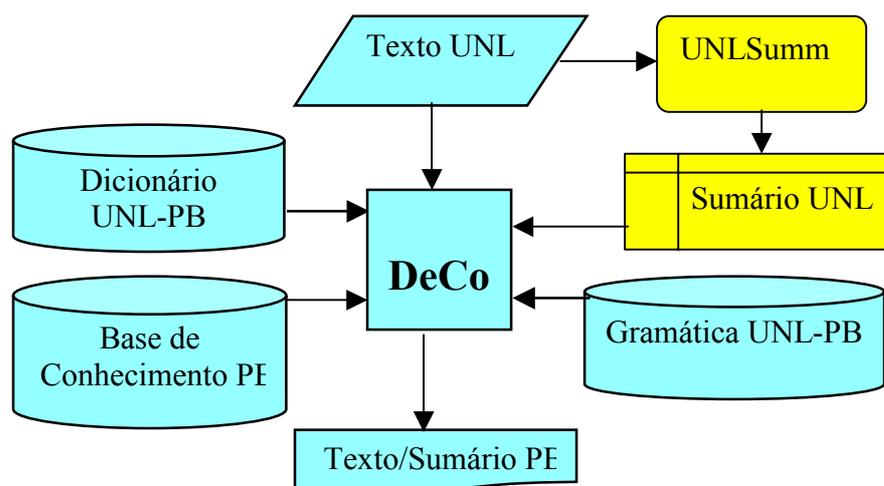


Figura 2: UNLSumm acoplado ao sistema de Decodificação UNL-Português

2.3 UNLSumm: O Mecanismo de Poda das Heurísticas de Sumarização

Com base nos RLs (*Relation Labels*) descritos na Seção 2.1, as heurísticas de poda a serem apresentadas aqui são responsáveis por identificar e excluir as informações menos relevantes de sentenças UNL. Como vimos, elas buscam relações binárias, ou RBs, que possam ser consideradas supérfluas e, logo, excluídas da sentença original. Sua identificação, por sua vez, é feita com base nos RLs, já que são esses os rótulos que remetem ao caráter semântico da relação. Assim, as heurísticas possuem base conceitual e foram inicialmente definidas a partir de análise de corpora de textos, conforme descrição em (Martins e Rino, 2002).

As heurísticas de sumarização somam atualmente, na versão 1.3, 84, divididas em dois grupos principais: Grupo A e Grupo B (Anexo 1). O Grupo A é composto por heurísticas que excluem apenas uma RB por vez. O Grupo B inclui heurísticas mais complexas, que excluem simultaneamente grupos de RBs inter-relacionadas. À poda de RBs dessa natureza chamamos poda encadeada.

Dentro de cada grupo, as heurísticas são novamente agrupadas de acordo com o RL focalizado, já que pode haver mais de uma forma possível de exclusão do mesmo caso conceitual. Assim, para uma melhor referência, as heurísticas são nomeadas HG.S.N, para G correspondendo ao seu grupo principal (A ou B); S indicando o subgrupo e N o número da heurística no subgrupo S. Por exemplo, as heurísticas que começam por HA.1 são relativas ao RL ‘met’ (*method*) no Grupo A. Portanto, a primeira heurística relativa a esse RL é nomeada HA.1.1, a segunda heurística é nomeada HA.1.2 e assim por diante (ver Anexo 1 para mais detalhes). Assim, quando nos referimos a um subgrupo de heurísticas neste relatório, estamos falando de um conjunto de heurísticas que focalizam um único RL.

3 O Estudo de Casos

O estudo de casos consistiu, basicamente, em gerar sumários utilizando o UNLSumm e, em seguida, avaliá-los. A avaliação foi realizada com base em duas métricas – textualidade e preservação do foco da sentença original, definidas abaixo.

3.1 Principais Métricas para Avaliação

3.1.1 Textualidade

Textualidade é a propriedade de um texto ser coerente e coeso (Rino, 1996). Assim, no nosso caso, considerando um texto mono-sentencial, uma sentença será considerada textual quando a ela puder ser associado um significado e quando ela apresentar ligações coesivas na superfície. Daqui em diante, vamos nos referir a uma sentença que apresenta tais características como sendo uma “sentença bem formada”.

3.1.2 Preservação do Foco

Esta medida consiste em avaliar se a idéia central da sentença fonte foi preservada, isto é, verificar se foi preservado o seu foco. Neste caso, dizemos que o sumário da sentença preserva o foco e, portanto, é um bom sumário, pois não omite o que é essencial. O foco de uma sentença ou sumário é a informação nova, em contrapartida à informação já conhecida. Fazemos correspondência, nesse caso, às noções respectivas de *focus* e *ground* de Vallduvi (1993). Por exemplo, considerando as sentenças ilustradas no Exemplo 1 como seqüenciais em um texto, podemos identificar, em [2], a ‘bola de João’ como *ground information*. Nessa sentença há ainda duas idéias centrais, correspondentes a dois *focci*: 1) a propriedade de a bola ser linda e 2) o fim a que ela se destina – o futebol.

[1]“João ganhou uma bola.” [2]“A linda bola é de futebol.”
--

Exemplo 1

Em nossas análises, consideramos que a informação principal (ou seja, a idéia central) da sentença coincidirá, na maioria das vezes, com a totalidade ou parte do *focus* da sentença. Portanto, para esse exemplo, se excluirmos as informações de que a bola ‘é linda e é de futebol’, não estaremos preservando a idéia central da sentença-fonte.

Em relação à textualidade, consideremos as sentenças no Exemplo 2. Se considerarmos que a informação *ground* correspondente a ‘toda a tarefa’ possa ser excluída da sentença 2 (o que seria verdadeiro, pois esta seria uma informação já compartilhada entre os interlocutores), o resultado correspondente poderia ser, após a decodificação UNL-Português, por exemplo, “Sua mãe ficou feliz ao ver que ele tinha feito.”. No entanto, essa frase não pode ser considerada textual, pois não transmite um significado, já que não apresenta o que foi que João fez. Embora isso possa ser recuperado pelo contexto da sentença 1, como estamos considerando a sumarização intra-sentencial, a sentença deve conseguir representar um significado isoladamente.

[1] “João fez toda sua tarefa.” [2] “Sua mãe ficou feliz ao ver que ele tinha feito toda a tarefa.”

Exemplo 2

As métricas de textualidade e preservação do foco são básicas em nosso estudo de casos: quando nos referimos a um sumário com qualidade satisfatória, entendemos que é um sumário que: a) é bem formado (e, logo, apresenta a propriedade de textualidade) e b) preserva o foco da sentença original. O uso dessas métricas é descrito a seguir.

3.1.3 Score dos Sumários

As duas métricas definidas acima foram usadas em conjunto, na avaliação dos sumários gerados pelo UNLSumm. Essa estratégia de avaliação é chamada de estratégia baseada em pontos de decisão (White et al., 2000). Em nossa proposta, tal pontuação não é uniforme: damos maior peso à preservação do foco, considerando que um sumário será útil se preservar a idéia central de sua fonte, mesmo que não seja satisfatoriamente bem formado. Para esses pontos de decisão, valem os *scores* apresentados na Tabela 2, a serem atribuídos a cada sumário UNL.

Textualidade do sumário UNL	Foco do sumário UNL	Score
Preservada	Preservado	6
Prejudicada	Preservado	5
Preservada	Preservado parcialmente	4
Prejudicada	Preservado parcialmente	3
Preservada	Ausente	2
Prejudicada	Ausente	1

Tabela 2: Medidas de avaliação de sumários UNL

Estabelecemos inicialmente que seriam consideradas aceitáveis as sentenças sumarizadas com *score* acima do ponto médio: 3,5. Assim, em nosso estudo de casos, sumários com *score* acima de 4,3 são considerados bons e, a partir de 5,1, muito bons. Além de atribuímos tais valores aos sumários sob análise, calculamos também sua taxa de compressão, definida a seguir.

3.1.4 Taxa de Compressão

A taxa de compressão de um sumário é medida pela diferença entre a quantidade de informação presente na sentença original e a que restou, no sumário (em porcentagem). Desse modo, essa taxa espelha o quanto foi *comprimido* ou *eliminado* do texto original, no processo de sumarização. No UNLSumm, essa taxa foi medida em relação ao número original e final de RBs das sentenças UNL envolvidas, i.e., se a sentença era originalmente formada por 16 RBs e, após a poda, resultou em um sumário UNL com 15 RBs, então a taxa de compressão foi de 6,25% ($1/16 * 100$).

3.2 Premissas e Hipóteses do Estudo de Casos

Nesta seção apresentamos as principais premissas do estudo de casos e as hipóteses a serem verificadas. As premissas se referem principalmente ao contexto do Projeto UNL,

devido ao fato de o UNLSumm utilizar as ferramentas desse projeto. As hipóteses são delineadas em função da verificação da qualidade dos sumários, visando, na verdade, verificar a qualidade e consistência das heurísticas de poda.

3.2.1 *Premissas*

Premissa 1 – No Projeto UNL os textos são representados sentença por sentença. Da mesma forma, o processamento para codificação/decodificação é realizado sentença por sentença. Na decodificação, supõe-se que o processamento seqüencial de sentenças do texto permite produzir textos bem formados. Similarmente, consideramos que sucessivas sumarizações intra-sentenciais resultam em sumários bem formados.

Premissa 2 – Embora não haja um consenso sobre o tamanho ideal de um sumário, normalmente considera-se que bons sumários mantêm de 5% a 30% do conteúdo do texto-fonte (Mani, 2001), ou seja, possuem taxas de compressão entre 70% e 95%. No entanto, para nosso tipo de poda intra-sentencial, essas taxas são muito altas, devido ao fato de não excluirmos sentença alguma do texto UNL fonte. Portanto, vamos, primeiramente, analisar as taxas de compressão obtidas unicamente no contexto de sumarização intra-sentencial, para então tentar estabelecer um intervalo de variação razoável para esse tipo de sumarização.

3.2.2 *Hipóteses a Serem Verificadas*

Na definição das hipóteses do estudo de casos, usamos vários conceitos, definidos no contexto da sumarização de textos UNL como seguem:

Aplicabilidade: essa noção expressa se um conjunto de heurísticas pode ser efetivamente aplicado a um corpus de textos UNL. Neste caso, uma heurística será aplicável se os casos conceituais que ela envolve ocorrerem nos corpora de investigação no estudo de casos.

Satisfatibilidade: o índice de satisfatibilidade de um resultado automático, i.e., de um sumário UNL, corresponde à medida correspondente às métricas de textualidade e preservação do foco definidas anteriormente..

Consistência: esta medida corresponde à capacidade do sistema de executar de forma correta e padronizada, ou seja, de aplicar com sucesso as heurísticas para qualquer tipo de sentença (i.e., de qualquer tamanho e sobre qualquer assunto), sem levar a conflitos na tomada de decisão sobre sua aplicação, para todos os casos previstos de dados de entrada.

Considerando esses conceitos, as hipóteses foram definidas como seguem:

Hipótese 1 – Aplicabilidade: As heurísticas definidas em função dos RLs são aplicáveis aos textos UNL em foco (i.e., àqueles dos corpora de investigação no estudo de casos). Dizer que uma heurística é aplicável implica afirmar que os casos conceituais que ela envolve são passíveis de ocorrência em códigos UNL fonte, dados como entrada do UNLSumm.

Hipótese 2 – Satisfatibilidade individual: A aplicação individual de heurísticas de qualquer grupo leva a sumários satisfatórios. Verificar se uma heurística é satisfatória implica, assim, verificar se há pelo menos um caso em que ela não leve a resultados negativos. De modo oposto, se para todos os casos possíveis de sumarização, a heurística

resulta em casos que não preservam o essencial, ela deverá ser excluída do rol de heurísticas do UNLSumm, por não privilegiar adequadamente os critérios de sumarização.

Hipótese 3 – Satisfatibilidade conjunta: A aplicação conjunta de heurísticas de qualquer grupo é satisfatória, i.e., permite sumarizar satisfatoriamente os textos UNL. Neste caso, verificar se a aplicação é satisfatória consiste em verificar se a satisfatibilidade é mantida quando se aplica mais de uma heurística a uma mesma sentença.

Hipótese 4 – Consistência: A aplicação conjunta de heurísticas é consistente, isto é, o UNLSumm consegue aplicar com sucesso as heurísticas a qualquer tipo de sentença, sem levar a conflitos na tomada de decisão.

Hipótese 5 – Interdependência: Há uma interdependência entre as heurísticas do Grupo A e do Grupo B, que determina sua ordem de aplicação. Essa interdependência indicará a estratégia de poda dos textos UNL, para alcançar as condições de satisfatibilidade.

3.3 Metodologia de Elaboração do Estudo de Casos

Para verificar as hipóteses acima, foram utilizados dois corpora de textos UNL: o Corpus UNU (gênero: legal, domínio: jurídico), totalizando 78 sentenças e 1301 RBs, e o Corpus Theses (gênero: científico, domínio: ciência da computação) totalizando 253 sentenças e 3035 RBs. O primeiro foi disponibilizado pelo Projeto UNL e o segundo foi manualmente codificado por nós a partir de textos científicos em Português⁴. Após a sumarização de cada sentença UNL, sua decodificação para o português foi feita manualmente para o Corpus UNU e automaticamente pelo DeCo para o Corpus Theses, como apresentado na Figura 2 (seção 2.2). para este corpus, cada sentença passou por um processo de pós-edição, de forma a eliminar erros provenientes do DeCo, que, por sua vez, está ainda em fase de desenvolvimento. Evitamos assim, avaliar erroneamente um sumário em função dos erros gerados pelo DeCo.

O estudo de casos contemplou somente uma sentença por vez, remetendo, assim, às noções de textualidade e preservação da idéia central delimitadas ao caso sentencial, conforme explicamos na Seção 3.1. Foram gerados diversos sumários, utilizando as diferentes heurísticas, de acordo com a hipótese a ser verificada. Consideramos seis casos:

Caso 1 – analisar a aplicabilidade de uma heurística por vez, do Grupo A (aplicação individual do Grupo A).

Caso 2 – analisar a aplicabilidade de uma heurística por vez, do Grupo B (aplicação individual do Grupo B).

Caso 3 – analisar a aplicabilidade de um conjunto de heurísticas do Grupo A (aplicação conjunta do Grupo A).

Caso 4 – analisar a aplicabilidade de um conjunto de heurísticas do Grupo B (aplicação conjunta do Grupo B).

Caso 5 – analisar a aplicabilidade de heurísticas do Grupo A, seguida da aplicação de heurísticas do Grupo B (aplicação conjunta ordenada pelo Grupo A).

Caso 6 – analisar a aplicabilidade de heurísticas do Grupo B, seguida da aplicação de heurísticas do Grupo A (aplicação conjunta ordenada pelo Grupo B).

⁴ Os textos em português do Corpus Theses são um subgrupo do Corpus DT, descrito em (Feltrim et al., 2001).

Para todos esses casos, temos o seguinte procedimento geral:

- Passo 1 (Busca) - Procuram-se as heurísticas aplicáveis à sentença UNL. Essas heurísticas são agrupadas no conjunto Hp (Heurísticas possíveis), percorrendo-se a sentença RB por RB. Para cada RB, todas as heurísticas do subgrupo relacionado ao seu RL são verificadas.
- Passo 2 (Ordenação) - As heurísticas do conjunto Hp são ordenadas⁵.
- Passo 3 (Aplicação) - A primeira heurística de Hp é aplicada e, em seguida, excluída da lista.
- Passo 4 (Atualização) - Verifica-se novamente todas as heurísticas de Hp, para ver se nenhuma ficou invalidada após o passo 2. Em caso afirmativo, a lista é atualizada, retirando-se as heurísticas inválidas.
- Passo 5 (Re-poda) – Se Hp ainda contiver heurísticas, voltar ao passo 2.

Vale ressaltar que o passo de atualização (passo 4) é necessário quando se aplica mais de uma heurística a uma mesma sentença, pois, como a aplicação de uma heurística muda a configuração do código UNL gradativamente, é possível que a lista de heurísticas aplicáveis (determinada no Passo 1) não seja mais realista após a primeira poda da sentença UNL. Esse passo, portanto, não é necessário nos casos 1 e 2, quando apenas uma heurística é aplicada a cada sentença. O mesmo ocorre, nesses casos, para o passo 5. Assim, para deixar claro quais passos são executados em cada caso, descrevemos a seguir cada um dos seis casos, em função da aplicação individual, da aplicação conjunta e da ordem de aplicação das heurísticas, respectivamente.

3.3.1 Casos 1 e 2

Estes casos consistiram apenas dos passos de busca (passo 1) e aplicação (passo 3), pois apenas uma única heurística foi aplicada a cada sentença por vez, sendo que o Caso 1 refere-se ao Grupo A e o Caso 2 ao Grupo B. Vale ressaltar que, quando a mesma heurística era aplicável a diferentes componentes de uma sentença, cada aplicação representou uma instância distinta de aplicação, sendo compilada, portanto, separadamente. Ou seja, mesmo que uma mesma heurística fosse aplicável várias vezes a uma mesma sentença, consideramos que cada aplicação resultou em um novo sumário.

3.3.2 Casos 3 e 4

Nesses casos consideramos a aplicação exaustiva de todas as heurísticas possíveis de um grupo a cada sentença. O Caso 3 refere-se a heurísticas do Grupo A e o Caso 4, a heurísticas do Grupo B. Assim, para a sumarização completa de uma sentença UNL, a fase de busca (passo 1) selecionou apenas as heurísticas do Grupo correspondente, não foi necessária a fase de ordenação e, após a fase de aplicação (passo 3), a lista Hp foi atualizada (passo 4) para, em seguida, continuar a poda (passo 5).

3.3.3 Casos 5 e 6

Estes são os casos que consideram aplicações conjuntas de heurísticas dos grupos A e B, porém, uma vez iniciando a poda por um grupo, segue-se exaustivamente buscando as heurísticas desse grupo, para depois passar às heurísticas do outro. Desse modo, não investigamos a aplicação intercalada de heurísticas de um e outro grupo, sendo que no

⁵ Este passo é necessário apenas nos casos onde são aplicadas heurísticas de ambos os grupos (5 e 6), como apresentado na seção 3.3.3.

Caso 5 iniciamos a poda pelas heurísticas do Grupo A e no Caso 6 pelas heurísticas do Grupo B. Assim, estes casos seguem todos os passos descritos no procedimento geral, distinguindo-se somente a forma como Hp é ordenada no passo 2. No Caso 5, as heurísticas do Grupo A são movidas para o início da lista Hp. No caso 6, as heurísticas do Grupo B são movidas para o início dessa lista.

No próximo capítulo apresentamos os resultados do estudo de casos descrito acima.

4 Resultados Obtidos com o Estudo de Casos

Neste capítulo apresentamos os resultados do processamento de cada corpus. Para cada caso, analisamos as médias gerais dos *scores* obtidos e, em seguida, analisamos separadamente as duas restrições fundamentais para nossos sumários, ou seja, a preservação do foco e a verificação da textualidade, contabilizando a porcentagem de sumários que apresentaram resultados satisfatórios para ambas as características. Como as heurísticas são agrupadas por RL dentro dos Grupos A e B, os resultados serão computados de acordo com esses subgrupos. Apresentamos também, juntamente com cada caso, um exemplo de uma sentença sumarizada pelo UNLSumm, discutindo a qualidade do resultado gerado automaticamente. A definição completa de cada heurística usada pode ser encontrada em (Martins e Rino, 2002) e é resumida também no Anexo 1.

4.1 Estudo de Casos para o Corpus UNU

4.1.1 Aplicação individual de heurísticas

4.1.1.1 Caso 1 – Heurísticas do Grupo A

A Tabela 3 apresenta um resumo dos resultados obtidos para a aplicação individual de heurísticas do Grupo A, i.e., aquelas que contemplam somente uma relação binária do código UNL de uma sentença. Como podemos observar, a média geral dos *scores* foi alta (5,61). Verificamos também que as heurísticas com o pior desempenho são as do subgrupo HA.7⁶, i.e., aquelas que contemplam o RL ‘mod’, que corresponde, em geral, a formações do tipo conceito-modificador, ou substantivo-adjetivo, no português. Analisando cada métrica separadamente (Tabela 4), verificamos que 82,20% dos sumários preservaram totalmente o foco do sumário, enquanto 72,03% foram considerados bem formados. Cada heurística do Grupo A comprime, em média, cerca de 7,71% da sentença-fonte.

Subgrupo	RL relacionado	Número de Heurísticas aplicadas	Média de compressão	Média dos scores
HA.2	man	10	7,3%	5,4
HA.3	plc	7	7,43%	6
HA.5	tim	38	8,5%	5,5
HA.6	pos	1	7%	6
HA.7	mod	167	6,57%	4,35
HA.10	qua	3	8%	6
HA.11	cnt	10	9,18%	6
Totais	7	236		
Médias			7,71%	5,61

Tabela 3: Estatística de poda para o Caso 1

Por fim, verificamos que as heurísticas de alguns subgrupos não foram aplicadas nenhuma vez, por exemplo, as heurísticas que contemplam o RL ‘met’. Analisando

⁶ Vide Anexo 1 para definição das heurísticas referenciadas por seu código neste texto, assim como Tabela 1, na Seção 2.

manualmente o corpus de textos, verificamos que em todos os casos de ocorrência do RL ‘met’, a informação não era independente, ou seja, só poderia ser excluída por uma heurística do Grupo B. Assim, os subgrupos HA.1 (RL ‘met), HA.4 (RL ‘aoj’), HA.8 (RL ‘nam’) e HA.9 (RL ‘src’) tiveram aplicabilidade 0 neste caso.

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	129	33	8	72,03%
Prejudicada	65	1	0	27,97%
Porcentagem	82,20%	14,41%	3,39%	

Tabela 4: Resultados para preservação do foco e textualidade no Caso 1

Exemplo de Aplicação de Heurísticas do Grupo A

A sentença [3] “*The ILO was established in 1919, to promote social justice for working people everywhere.*”, cuja codificação UNL é apresentada no Exemplo 3, foi resumida pela heurística HA.3.3, cujo subgrupo contempla a relação semântica *place*, rotulada pelo RL ‘plc’. O sumário UNL resultante é apresentado no Exemplo 4 e poderia ser decodificado como “*The ILO was established in 1919, to promote social justice for working people.*”. Como podemos observar, a sentença continua apresentando elos coesivos entre os componentes sentenciais, fazendo com que tal sumário seja bem formado (ou textual). Além disso, considerando que o foco da sentença original seja o propósito da existência da organização ILO, ele também preserva a informação principal da sentença—fonte e, portanto, cumpre sua função principal, segundo nossas métricas essenciais de sumarização. Sumários como esse foram avaliados com o *score* máximo 6.

```
[S]
obj(establish.@past,ILO.@def)
tim(establish.@past,1919)
pur(establish.@past,promote)
obj(promote,justice)
mod(justice,social)
pur(justice,people)
mod(people,working)
plc(people,everywhere)
[/S]
```

Exemplo 3: Codificação UNL de [3]

```
[S]
obj(establish.@past,ILO.@def)
tim(establish.@past,1919)
pur(establish.@past,promote)
obj(promote,justice)
mod(justice,social)
pur(justice,people)
mod(people,working)
[/S]
```

Exemplo 4: Possível sumário UNL de [3]

4.1.1.2 Caso 2 – Heurísticas do Grupo B

A Tabela 5 resume os resultados para a aplicação individual de heurísticas do Grupo B, i.e., aquelas que contemplam a poda de grupos de relações binárias do código UNL de uma sentença, porém, a partir de um RL predominante, que é o RL indicado nessa tabela. Considerando as médias gerais dos *scores* obtidos, a maioria das podas dos subgrupos assinalados é aceitável, com uma média geral de 4,43. Entretanto, os subgrupos HB.4 e HB.14, que disparam a poda a partir de RBs que envolvem, respectivamente, as relações semânticas *attribute of an object* (RL ‘aoj’) e *purpose* (RL ‘pur’), obtiveram resultados inaceitáveis. Voltaremos a essa questão no Capítulo 5, para discutir melhorias das estratégias de sumarização. Analisando cada métrica separadamente (Tabela 6),

verificamos que 55,41% dos sumários preservaram totalmente o foco do sumário, enquanto 80,25% foram considerados bem formados. Cada heurística do Grupo B comprime, em média, cerca de 37,23% da sentença-fonte.

Por fim, verificamos que as heurísticas de vários subgrupos não foram aplicadas nenhuma vez. Os subgrupos que tiveram aplicabilidade 0 foram: HB.9 (RL 'src'), HB.13 (RL 'rsn'), HB.15 (RL 'dur'), HB.16 (RL 'con'), HB.17 (RL 'agt') e HB.18 (RL 'ben')⁷. Para investigar o porquê da não aplicabilidade desses subgrupos, analisamos manualmente o corpus de textos.

Para o subgrupo HB.9, verificamos que houve apenas uma ocorrência do RL 'src' em todo o corpus. Na única sentença UNL onde esse RL aparece, a heurística não foi aplicada porque, das 10 RBs da sentença, 9 se referiam a informações apontadas pelo RL 'src'. Ou seja, nesse caso a informação indicada pelo RL era essencial e foi corretamente mantida.

Já os subgrupos HB.13, HB.15, HB.16 e HB.18 tiveram aplicabilidade 0 por não haver nenhuma ocorrência dos RLs 'rsn', 'dur', 'con' e 'ben' em todo o corpus em estudo (Corpus UNU).

Por fim, em relação ao subgrupo HB.17, verificamos que houve ocorrências do RL 'agt', mas nenhuma que satisfizesse as condições das heurísticas desse subgrupo.

Subgrupo	RL relacionado	Número de Heurísticas aplicadas	Média de compressão	Média dos scores
HB.1	met	8	40%	5,38
HB.2	man	6	37,33%	4,67
HB.3	plc	19	20,89%	5,47
HB.4	aoj	7	69,19%	2,57
HB.5	tim	3	16,67%	5
HB.7	mod	72	23,14%	4,82
HB.12	scn	6	42,11%	4,5
HB.14	pur	36	48,52%	3
Totais	8	157		
Médias			37,23%	4,43

Tabela 5: Estatísticas de poda para o Caso 2

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	65	21	40	80,25%
Prejudicada	22	6	3	19,75%
Porcentagem	55,41%	17,20%	27,39%	

Tabela 6: Resultados para preservação do foco e textualidade no Caso 2

Exemplo de Aplicação de Heurísticas do Grupo B

Como o subgrupo HB.4 apresentou a pior média dos *scores*, vamos apresentar aqui um exemplo de uma sentença sumarizada por uma heurística desse subgrupo, discutindo as causas desse mau desempenho.

⁷ Alguns RLs são utilizados apenas no Grupo A ou B. Assim, não existem subgrupos HB.6, HB.8, HB.10 e HB.11. Vide no Anexo1 quais RLs são contemplados em cada grupo de heurísticas.

A sentença [4] “*The International Network on Water, Environment and Health (UNU/INWEH) operating since 1996 from Hamilton, Canada, is an interdisciplinary and global network of research and training on water pollution and management.*”, cuja codificação UNL é apresentada no Exemplo 5, foi sumarizada por meio da heurística HB.4.3, que contempla atributos de objetos (RL ‘aoj’) e demais relações binárias contextualizadas a partir de tal conceito. O sumário UNL correspondente é apresentado no Exemplo 6, podendo ser decodificado como “*The International Network on Water, Environment and Health (UNU/INWEH) is operating since 1996 from Hamilton, Canadá.*”. Pode-se considerar que, neste exemplo, a informação principal não foi mantida. No entanto, o sumário continua sendo uma sentença bem formada. Assim, o sumário recebeu o *score* 2.

Observando a sentença-fonte UNL, pode-se considerar que este desempenho ruim da heurística deve-se à ambigüidade do RL ‘aoj’. Ao mesmo tempo em que ele pode ser utilizado para representar atributos (geralmente adjetivos) supérfluos, em alguns casos, como neste exemplo, ele é utilizado para relacionar o conceito principal da sentença (*The International Network on Water, Environment and Health*) ao seu predicado (*an interdisciplinary and global network*) e, portanto, é uma relação binária essencial.

```
[S]
aoj(network.@present.@indef,International Network on Water Environment and Health.@def)
cnt(International Network on Water Environment and Health.@def, UNU/INWEH)
mod(International Network on Water Environment and Health.@def,:01)
tmf:01(operate, 1996)
plf:01(operate, Hamilton)
mod:01(Hamilton, Canada)
mod(network.@present.@indef,global)
and(global, interdisciplinary)
mod(network.@present.@indef,:02)
and:02(training, research)
scn(:02,:03)
and:03(management, pollution)
mod:03(pollution, water)
[/S]
```

Exemplo 5: Codificação UNL de [4]

```
[S]
cnt(International Network on Water Environment and Health.@def, UNU/INWEH)
mod(International Network on Water Environment and Health.@def,:01)
tmf:01(operate, 1996)
plf:01(operate, Hamilton(icl>city))
mod:01(Hamilton, Canada)
[/S]
```

Exemplo 6: Possível sumário UNL de [4]

4.1.1.3 Crítica Parcial

Os casos 1 e 2 podem ser analisados em relação às hipóteses 1 e 2 descritas anteriormente. Assim, temos o seguinte:

1) A maioria das heurísticas de ambos os grupos é aplicável. No entanto, houve casos de subgrupos de heurísticas com aplicabilidade 0, devido às características do corpus utilizado.

2) Todas as heurísticas aplicadas apresentaram resultados satisfatórios, i.e., cada heurística produziu pelo menos um sumário satisfatório. No entanto, os subgrupos de heurísticas HB.4 (RL ‘aoj’) e HB.14 (RL ‘pur’) obtiveram uma média dos *scores* insatisfatória, abaixo de 3,5. Verificamos também que a satisfatibilidade das heurísticas do Grupo A é maior do que as do Grupo B.

4.1.2 Aplicação conjunta de heurísticas

4.1.2.1 Caso 3 – Heurísticas do Grupo A

A Tabela 7 mostra alguns exemplos de resultados da aplicação conjunta de heurísticas do Grupo A a 73 sentenças do Corpus UNU (nas demais sentenças do corpus nenhuma heurística foi considerada aplicável). As heurísticas são apresentadas nessa tabela em sua ordem de aplicação. É importante notar que diferentes ordens podem levar a diferentes resultados e, logo, a diferentes sumários, devido à atualização intermediária da lista de heurísticas aplicáveis. Em termos conceituais, isso corresponde a alterar o contexto do código UNL a cada poda de uma RB e, portanto, alterar a possível relevância dos segmentos restantes. Avaliando-se os resultados da aplicação conjunta de heurísticas do Grupo A a todas as sentenças do Corpus UNU, obtivemos uma média de avaliação de 4,3. Analisando cada métrica separadamente (Tabela 8), verificamos que pouco mais da metade dos sumários preservou o foco e resultou em uma sentença bem formada, para uma taxa de compressão média de 26,36%.

Sentença	Heurísticas aplicadas	Compressão	Score
Admin-f	HA.7.5, HA.7.5	15%	6
Biolac-f	HA.11.1, HA.5.3, HA.3.3	30%	6
Board-f	HA.7.7, HA.10.1, HA.7.5, HA.7.5, HA.6.1, HA.7.6	46%	5
Condir-f	HA.11.1, HA.7.5	11%	6
Direct-f	HA.7.6, HA.7.6, HA.7.5	25%	4
Eca-f	HA.5.6, HA.7.7, HA.7.6.	20%	6
Ece-f	HA.5.3, HA.7.6, HA.7.7, HA.7.6, HA.7.7, HA.2.2, HA.7.7, HA.7.5	38%	2
...
Médias		26,36%	4,3

Tabela 7: Estatísticas de poda parciais para o Caso 3⁸

Foco Textualidade	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Preservada	18	15	6	53,42%
Prejudicada	19	12	3	46,58%
Porcentagem	50,68%	36,99%	12,33%	

Tabela 8: Resultados para preservação do foco e textualidade no Caso 3

Exemplo da Aplicação Conjunta de Heurísticas do Grupo A

A sentença [5] “UNU Programme for Biotechnology in Latin America and the Caribbean (BIOLAC)), established in 1998, in Caracas, Venezuela, promotes research and training in biotechnology in the Region.”, referenciada por Biolac-f na Tabela 6, foi

⁸ A coluna *Sentença* indica por um nome fantasia cada sentença do Corpus UNU.

sumarizada por três heurísticas (vide coluna ‘Heurísticas aplicadas’). Sua codificação UNL e correspondente sumário são apresentados nos Exemplos 7 e 8, respectivamente. Esse sumário UNL pode ser decodificado como “*UNU Programme for Biotechnology in Latin America and the Caribbean, established in Caracas, Venezuela, promotes research and training in biotechnology.*”. Esta sentença é bem formada, pois apresenta elos coesivos e é coerente, permitindo, assim, a veiculação de um significado. Além disso, ela exclui somente informações periféricas, quando comparada à sentença original. Logo, o sumário foi avaliado com o *score* máximo 6.

```
[S]
agt(promote.@present, UNU Programme for Biotechnology in Latin America and the
  Carribbean)
cnt(UNU Programme for Biotechnology in Latin America and the Carribbean, BIOLAC)
obj(establish.@past, UNU Programme for Biotechnology in Latin America and the Carribbean)
tim(establish.@past, 1998)
plc(establish.@past, Caracas)
mod(Caracas, Venezuela)
obj(promote.@present,:02)
and:02(training, research)
scn(:02, biotechnology)
plc(biotechnology, region.@def)
[/S]
```

Exemplo 7: Codificação UNL de [5]

```
[S]
agt(promote.@present, UNU Programme for Biotechnology in Latin America and the
  Carribbean)
obj(establish.@past, UNU Programme for Biotechnology in Latin America and the Carribbean)
plc(establish.@past, Caracas)
mod(Caracas, Venezuela)
obj(promote.@present, :02)
and:02(training, research)
scn(:02, biotechnology)
[/S]
```

Exemplo 8: Possível sumário UNL de [5]

4.1.2.2 Caso 4 – Heurísticas do Grupo B

A Tabela 9 resume os resultados para a aplicação conjunta de heurísticas do Grupo B a 67 sentenças do Corpus UNU (novamente, as sentenças restantes desse corpus não foram sumarizadas pois nenhuma heurística do Grupo B foi considerada aplicável), totalizando uma média de avaliação de 3,58. Observamos também que, na maioria dos casos em que a avaliação foi abaixo de 3,5, as podas se deveram a heurísticas dos subgrupos HB.7, HB.4 e HB.14, correspondentes aos RLs ‘mod’, ‘aoj’ e ‘pur’, respectivamente. Apesar desse resultado, a maioria dos sumários foi considerada bem formada, apresentando um índice de textualidade de 70,15% (Tabela 10). Considerando que apenas 37,31% deles preservaram o foco da sentença original, a baixa avaliação geral dos resultados parece ser devida à não preservação do foco. Verificamos, ainda, que há uma relação entre a taxa de compressão e o índice de avaliação dos sumários: quanto maior essa taxa, menor o *score* da sentença correspondente. A média geral das taxas de compressão das sentenças do corpus foi de 48,03%, praticamente o dobro da taxa de compressão média do caso anterior.

Sentença	Heurísticas aplicadas	Compressão	Score
Admin-f	HB.7.4	23%	6
Biolac-f	HB.3.5, HB.12.1	40%	5
Board-f	HB.1.4	38%	6
Condir-f	HB.7.4, HB.14.3	70%	1
Direct-f	HB.7.4, HB.7.3	58%	4
Eca-f	HB.7.3, HB.7.3, HB.14.2	53%	6
Rector-f	HB.7.4, HB.7.3	61%	3
...
Médias		48,03%	3,58

Tabela 9: Estatísticas de poda parciais para o Caso 4.

	Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade					
Preservada		16	10	21	70,15%
Prejudicada		9	3	8	29,85%
Porcentagem		37,31%	19,40%	43,28%	

Tabela 10: Resultados para preservação do foco e textualidade no Caso 4

Exemplo da Aplicação Conjunta de Heurísticas do Grupo B

A sentença [6] “*The Rector is the Chief Academic and Administrative Officer of the University, with the responsibility for the direction, organization and administration of its overall programme.*”, referenciada na Tabela 9 por *Rector-f*, cujo código UNL é ilustrado no Exemplo 9, foi sumarizada pelas heurísticas HB.7.4 e HB.7.3, ambas correspondentes à relação semântica *modifier* (RL ‘mod’) – modificador de um conceito. O sumário UNL resultante é apresentado no Exemplo 10 e pode ser decodificado como “*The Rector is the Chief Officer of the University, with the responsibility.*”.

```
[S]
aoj(:01.@present,rector.@def)
mod:01(officer.@def, chief)
mod:01(officer.@def, administrative)
and:01(administrative,academic)
mod:01(officer.@def, university.@def)
mod(:01.@present, :02)
obj:02(with, responsibility.@def)
mod:02(responsibility.@def,:03)
and:03(administration, organization)
and:03(organization, direction.@def)
mod:02(:03,programme)
mod:02(programme, it)
mod:02(programme, overall)
[/S]
```

Exemplo 9: Codificação UNL de [6]

```
[S]
aoj(:01.@present,rector.@def)
mod:01(officer.@def, chief)
mod:01(officer.@def, university.@def)
mod(:01.@present, :02)
obj:02(with, responsibility.@def)
[/S]
```

Exemplo 10: Possível sumário UNL de [6]

Claramente, a textualidade do sumário apresentado foi prejudicada, devido à incompletude do sintagma preposicional. Mesmo que a decodificação resultasse em outras escolhas lingüísticas, isto teria ocorrido, devido ao fato de termos segmentos UNL *soltos* (*dangling segments*) no sumário UNL, i.e., sem conexão com o restante do código, o que

implica dizer que as realizações lingüísticas correspondentes serão mal formadas (e, logo, não textuais). Além disso, podemos considerar que a informação principal foi apenas parcialmente preservada, pois o sumário omite as funções desempenhadas pelo reitor, as quais, no contexto atual, seriam relevantes. Sumários com problemas dessa natureza não foram avaliados com *scores* superiores a 3.

4.1.2.3 Crítica Parcial da Aplicação Conjunta de Heurísticas dos Grupos A e B

Os casos 3 e 4 referem-se às hipóteses 1, 3 e 4 descritas na seção 3.2. Assim, temos o seguinte:

- 1) *Aplicabilidade* - As heurísticas mantêm sua aplicabilidade quando aplicadas em conjunto a uma sentença. Verificamos também que todos os subgrupos de heurísticas aplicados nos casos 1 e 2 mantiveram sua aplicabilidade nestes casos, indicando que a aplicação conjunta não afeta a aplicabilidade das heurísticas.
- 2) *Satisfatibilidade conjunta* - A aplicação conjunta tanto de heurísticas do Grupo A quanto do Grupo B é satisfatória. Verificamos também que a satisfatibilidade é maior aplicando-se heurísticas do Grupo A, o que sugere que, ao se optar por apenas um grupo de heurísticas, o Grupo A deve ser priorizado, pois tende a garantir sumários mais satisfatórios. Esta nova hipótese será discutida novamente no capítulo 5.
- 3) *Consistência* - A aplicação conjunta de heurísticas foi consistente para todos os dados de entrada.

4.1.3 Aplicação de Heurísticas Segundo sua Ordem

Os Casos 5 e 6 tiveram dois objetivos principais: 1) verificar a importância da ordem de aplicação das heurísticas de acordo com o grupo (Hipóteses 1 e 4) e b) verificar a interdependência da aplicação conjunta de heurísticas de ambos os grupos (Hipótese 5).

4.1.3.1 Caso 5 – Aplicação de Heurísticas do Grupo A, depois do Grupo B

A Tabela 11 resume os valores obtidos ao se privilegiar a poda por heurísticas do Grupo A, seguidas da poda por heurísticas do Grupo B, mostrando os valores de algumas das 73 sentenças sumarizadas neste caso. Embora a média geral de avaliação dos sumários tenha sido aceitável (3,56), houve 38 casos com *score* abaixo de 3,5. Além disso, menos da metade dos sumários preservou o foco (34,25%) e se apresentou bem formada (42,47%) (Tabela 12). Neste caso, a média das taxas de compressão dos sumários foi de 40,41%.

Sentença	Heu. Grupo A	Heu. Grupo B	Total	Compressão	Score
Admin-f	2	1	3	69%	3
Biolac-f	3	1	4	50%	6
Board-f	6	0	6	46%	5
Condir-f	2	2	4	35%	5
Direct-f	3	1	4	41%	5
Eca-f	3	1	4	33%	6
Escwa-f	3	1	4	45%	2
...
Totais	236	45	281		
Médias				40,41%	3,56

Tabela 11⁹: Estatísticas de poda parciais para o Caso 5

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	8	10	13	42,47%
Prejudicada	17	18	7	57,53%
Porcentagem	34,25%	38,36%	27,40%	

Tabela 12: Resultados para preservação do foco e textualidade no Caso 5

Exemplo da Aplicação de Heurísticas do Grupo A, depois do Grupo B

A sentença [7] “*ESCWA was established in 1973 to promote economic cooperation and integration and aid economic and social development in Western Asia.*” (referenciada como *Escwa-f* na Tabela 11), cuja codificação UNL é apresentada no Exemplo 11, foi resumida pelas heurísticas HA.5.3, HA.7.7, HA.3.3 e HB.7.4, correspondentes às relações semânticas de *time* (RL ‘tim’), *modifier* (RL ‘mod’), *place* (RL ‘plc’) e novamente *modifier*. O sumário UNL resultante é apresentado no Exemplo 12, com possível decodificação dada por “*ESCWA was established to promote cooperation and integration and aid development.*”.

Analisando a sentença original, a sentença resumida pode ser considerada bem formada, mas não preserva sua idéia principal, por excluir as especificidades do tipo de ajuda prestado pela ESCWA (de caráter econômico e social, limitado à Ásia Ocidental), tendo sido avaliada, portanto, com o *score* 2. A exclusão dessas informações foi feita pelas heurísticas HA.7.7 e HB.7.4, relacionadas justamente à relação semântica *modifier* (RL ‘mod’).

⁹ Nesta tabela, as colunas “Heu. Grupo A” e “Heu. Grupo B” indicam quantas heurísticas do Grupo A e B, respectivamente, foram aplicadas à sentença. A coluna “Total” indica o número total de heurísticas aplicadas.

```
[S]
obj(establish.@past,ESCWA)
tim(establish.@past, 1973)
pur(establish.@past, :01)
and:01(aid, promote)
obj:01(promote, :02)
and:02(integration, cooperation)
mod:01(:02,economic:01)
obj:01(aid, development)
mod:01(development, social)
and:01(social, economic:02)
plc(:01,Western Asia)
[/S]
```

Exemplo 11: Codificação UNL de [7]

```
[S]
obj(establish.@past,ESCWA)
pur(establish.@past, :01)
and:01(aid, promote)
obj:01(promote, :02)
and:02(integration, cooperation)
obj:01(aid, development)
[/S]
```

Exemplo 12: Possível sumário UNL de [7]

4.1.3.2 Caso 6 – Aplicação de Heurísticas do Grupo B, depois do Grupo A

A Tabela 13 resume os valores obtidos ao se privilegiar a poda por heurísticas do Grupo B, seguidas da poda por heurísticas do Grupo A. A média geral de avaliação dos sumários é inaceitável (2,85), enquanto a média de compressão resultou bastante alta em relação aos outros casos (56,06%). Observamos também que, de um modo geral, os piores resultados correspondem às maiores taxas de compressão. Verificamos, ainda, que a maioria dos sumários é bem formada (índice de textualidade de 58,90%), mas não preserva o foco (Tabela 14). Na verdade, este não foi preservado nem total, nem parcialmente, em 58,90% dos casos.

Sentença	Grupo A	Grupo B	Total	Compressão	Score
Admin-f	0	1	1	23%	6
Biolac-f	2	2	4	60%	5
Board-f	3	1	4	61%	5
Condir-f	1	2	3	76%	1
Direct-f	1	2	3	66%	2
Eca-f	1	3	4	60%	6
Ece-f	0	1	1	80%	1
...
Totais	8	92	185		
Médias				56,06%	2,85

Tabela 13: Estatísticas de poda parciais para o Caso 6

Foco Textualidade	Foco			Porcentagem
	Preservado	Preservado Parcialmente	Ausente	
Preservada	8	9	26	58,90%
Prejudicada	8	5	17	41,10%
Porcentagem	21,92%	19,18%	58,90%	

Tabela 14: Resultados para preservação do foco e textualidade no Caso 6

Exemplo da Aplicação de Heurísticas do Grupo B, depois do Grupo A

Um resultado ruim obtido neste caso também se deve à sumarização da sentença [5] - *biolac-f* – apresentada no Caso 3 e reproduzida aqui: “*UNU Programme for Biotechnology in Latin America and the Caribbean (BIOLAC)*), established in 1998, in

Caracas, Venezuela, promotes research and training in biotechnology in the Region.”. Neste Caso, ela foi sumarizada pelas heurísticas HB.3.5, HB.12.1, HA.11.1 e HA.5.3, que correspondem às relações semânticas *place* (RL ‘plc’), *scenario* (RL ‘scn’), *content* (RL ‘cnt’) e *time* (RL ‘tim’). O sumário UNL correspondente é apresentado no Exemplo 13, podendo ser decodificado como “*UNU Programme for Biotechnology in Latin America and the Caribbean, established, promotes research and training.*”.

```
[S]
agt(promote.@present, UNU Programme for Biotechnology in Latin America and the
  Carribbean)
obj(establish.@past, UNU Programme for Biotechnology in Latin America and the
  Carribbean)
plc(establish.@past, Caracas)
mod(Caracas, Venezuela)
obj(promote.@present, :02)
and:02(training, research)
scn(:02, biotechnology)
[/S]
```

Exemplo 13: Outro possível sumário UNL para [5]

Muito embora esse sumário não seja bem formado, pois a introdução da palavra “*established*” prejudica sua coerência, devido à falta de conectividade com o restante da sentença, atribuímos a ele o *score* 5. Este valor se deve unicamente à preservação da idéia central da sentença-fonte: é possível identificar a área de atuação do órgão (biotecnologia), assim como sua forma de atuação (a promoção de pesquisa e treinamento). Além disso, o local mencionado originalmente é preservado no sumário.

4.1.3.3 Crítica Parcial Sobre a Ordem de Aplicação de Heurísticas dos Grupos A e B

Os casos 5 e 6 referem-se às hipóteses 1, 4 e 5 descritas na seção 3.2. Assim, temos o seguinte:

- 1) *Aplicabilidade* - As heurísticas são aplicáveis independentemente da ordem de aplicação.
- 2) *Consistência* - A aplicação de heurística foi consistente para todos os dados de entrada, independentemente da ordem de aplicação das heurísticas.
- 3) *Interdependência* - Verificamos que realmente há uma interdependência entre as heurísticas do Grupo A e do Grupo B, pois os resultados obtidos nos casos 5 e 6 foram significativamente diferentes. Os dados indicam que, devido a essa interdependência, a melhor estratégia para poda é aplicar primeiro todas as heurísticas do Grupo A possíveis e, depois, as do Grupo B.

4.1.4 Crítica Geral Sobre o Corpus UNU

Após a análise parcial de cada caso, podemos discutir os resultados do estudo realizado com o Corpus UNU, de acordo com as cinco hipóteses adotadas.

1) Aplicabilidade - A maioria das heurísticas de ambos os grupos é aplicável. No entanto, os subgrupos HA.1 (RL ‘met’), HA.4 (RL ‘aoj’), HA.8 (RL ‘nam’), HA.9 (RL ‘src’), HB.13 (RL ‘rsn’), HB.15 (RL ‘dur’), HB.16 (RL ‘con’), HB.17 (RL ‘agt’) e HB.18 (RL ‘ben’) obtiveram aplicabilidade 0 neste corpus. Esse resultado pode tanto indicar que o corpus é muito restrito ou que as heurísticas têm pouca ou nenhuma aplicabilidade e, portanto, não serão úteis para a sumarização. Pretendemos verificar essas alternativas ao

repetir o estudo de casos para o Corpus Theses, descrito em 4.2. Verificamos ainda que a aplicabilidade é mantida tanto na aplicação conjunta de heurísticas quanto considerando qualquer ordem de aplicação.

2) Satisfatibilidade individual - Todos os subgrupos de heurísticas apresentaram resultados satisfatórios. No entanto, a maioria dos sumários insatisfatórios foram gerados por heurísticas dos subgrupos HB.4 (RL ‘aoj’), HB.7 (RL ‘mod’) e HB.14 (RL ‘pur’). Para investigar o motivo desses resultados, analisamos manualmente as sentenças sumarizadas por essas heurísticas. Para o primeiro subgrupo, verificamos que o RL ‘aoj’ é utilizado para representar diferentes tipos de informações, sendo esse possivelmente o motivo pelo qual as heurísticas relacionadas a esse RL levam a resultados bons em alguns casos mas ruins em outros, como nos exemplos 5 e 6 na seção 4.1.1.2. O mesmo ocorre com o subgrupo HB.7, pois em algumas sentenças o RL ‘mod’ representa informações irrelevantes e, em outras, informações essenciais. Ainda em relação a este subgrupo, verificamos que em diversos casos ele foi utilizado de forma inconsistente com a própria definição, representando informações que deveriam ser representadas por outros RLs. Em relação ao subgrupo HB.14, o resultado ruim parece ser devido às características do corpus, no qual a maioria das sentenças indica o propósito de alguma organização da ONU. Assim, como o RL ‘pur’ representa a relação semântica *purpose*, ou propósito, na maioria dos casos uma informação importante foi excluída. Por exemplo, a sentença [8] “*Created in 1948, ECLAC has a mission to coordinate policies and promote economic development in the Latin American and Caribbean regions.*”, cuja codificação UNL é apresentada no Exemplo 14, foi sumarizada no Caso 2 pela heurística HB.14.3, gerando o sumário UNL apresentado no Exemplo 15, correspondente a “*Created in 1948, ECLAC has a mission*”. Como podemos observar, esse sumário não preserva a idéia central da sentença fonte, pois não informa a missão da organização ECLAC. No entanto, a sentença é bem formada e, portanto, recebeu o *score* 2. A influência do corpus nesse resultado será também verificada na próxima seção, ao realizarmos o estudo com o Corpus Theses. Ainda em relação à satisfatibilidade individual, verificamos que as heurísticas do Grupo A apresentam resultados melhores do que as do Grupo B.

```
[S]
aoj(have.@entry,ECLAC)
obj(create.@past,ECLAC)
tim(create.@past,1948)
obj(have.@entry,mission.@indef)
pur(mission.@indef,:01)
and:01(promote.@entry,coordinate)
obj:01(coordinate,policy.@pl)
obj:01(promote.@entry,development)
mod:01(development,economic)
plc(:01,region.@def.@pl)
mod(region.@def.@pl,caribbean)
and(caribbean,latin american)
[/S]
```

Exemplo 14: Codificação UNL de [8]

```
[S]
aoj(have.@entry,ECLAC)
obj(create.@past,ECLAC)
tim(create.@past,1948)
obj(have.@entry,mission.@indef)
[/S]
```

Exemplo 15: Possível sumário UNL de [8]

3) Satisfatibilidade conjunta - A aplicação conjunta tanto de heurísticas do Grupo A quanto do Grupo B é satisfatória, pois a maioria dos sumários gerados nos casos 3 e 4 foi avaliada com *score* acima de 3,5. Verificamos também que a satisfatibilidade é maior aplicando-se heurísticas do Grupo A. Além disso, verificamos que o desempenho inferior

do Grupo B ocorreu devido a falhas na preservação da idéia central. Como observamos pelas Tabelas 4 e 6, as heurísticas do Grupo B levam a melhores resultados na textualidade, que é preservada em 80,25% dos sumários produzidos por heurísticas do Grupo B e em 72,03% dos sumários produzidos por heurísticas do Grupo A. No entanto, no Grupo B apenas 55,41% dos sumários preservam o foco, sendo que 82,20% dos sumários preservam o foco no Grupo A. Assim, como consideramos a preservação da idéia central mais importante do que a preservação da textualidade, podemos considerar que as heurísticas do Grupo A tiveram um desempenho melhor, o que sugere que, ao se optar por apenas um grupo de heurísticas, o Grupo A deve ser priorizado, pois tende a garantir sumários mais satisfatórios. Verificamos também, por meio dos casos 5 e 6, que a satisfatibilidade é mantida qualquer que seja a ordem de aplicação de heurísticas.

4) Consistência - A aplicação conjunta de heurísticas foi consistente para todos os dados de entrada, tanto utilizando qualquer um dos grupos de heurísticas (casos 3 e 4), quanto adotando qualquer ordem de aplicação (casos 5 e 6).

5) Interdependência - Verificamos que realmente há uma interdependência entre as heurísticas do Grupo A e do Grupo B, pois os resultados obtidos nos casos 5 e 6 foram significativamente diferentes, sendo que ao se aplicar primeiro heurísticas do Grupo A os resultados são melhores. Esse resultado se deve, possivelmente, ao fato de ser aplicado um número maior de heurísticas do Grupo A do que do Grupo B e, como observamos nas hipóteses 2 e 3, as heurísticas do Grupo A possuem uma maior satisfatibilidade. Devido a essa interdependência, os resultados indicam que uma estratégia melhor para poda pode ser sempre aplicar primeiro as heurísticas do Grupo A.

Por fim, uma observação importante se refere à dependência entre a taxa de compressão e a qualidade do sumário. Como indicam a Tabela 15 e o Gráfico 1, quanto maior a taxa de compressão, pior o resultado. Isso sugere que, quanto mais RBs forem excluídas, maior a probabilidade de se excluir informações relevantes do texto-fonte. Como as heurísticas do Grupo B sempre levam a maiores taxas de compressão em relação aos casos similares envolvendo o Grupo A, essas taxas sugerem que o desempenho dessas heurísticas será inferior àquele das heurísticas do Grupo A. Ainda, o Gráfico 1 sugere que sumários com taxa de compressão a partir de 41% tendem a gerar resultados insatisfatórios, com sumários avaliados com *scores* abaixo de 3,5. Consequentemente, esses resultados sugerem que as melhores abordagens são as utilidades nos casos 1, 2, 3 e 5, pois, nesses casos, a média das taxas de compressão se situa abaixo de 41%.

	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6
Média dos Scores	5,6	4,43	4,3	3,58	3,58	2,85
Compressão	7,71%	37,23%	26,36%	48,03%	40,41%	56,06%

Tabela 15: Estatística de poda global para o Corpus UNU

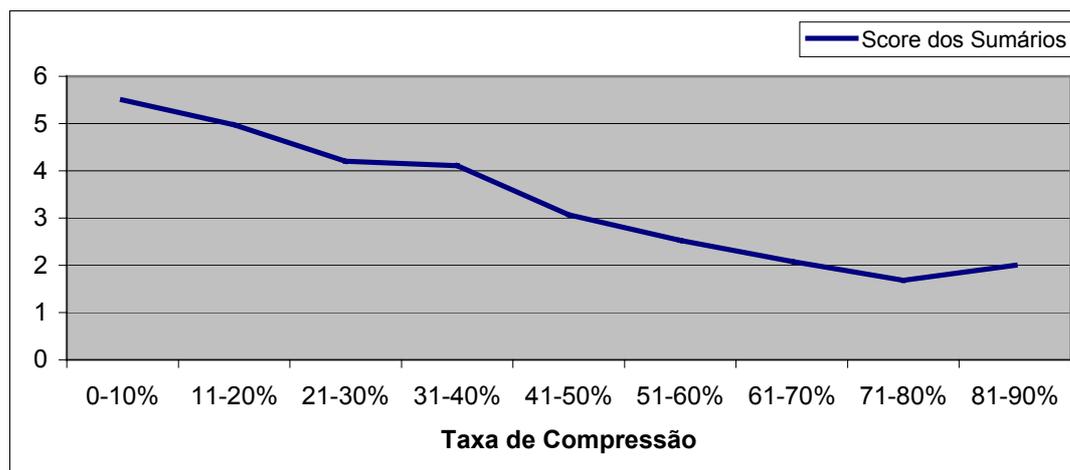


Gráfico 1: Desempenho dos sumários do Corpus UNU em relação à taxa de compressão

Para verificarmos a possível influência, nos resultados, do corpus utilizado, repetimos o estudo de casos para o Corpus Theses, descrito na próxima seção.

4.2 Estudo de Casos para o Corpus Theses

4.2.1 Aplicação individual de heurísticas

4.2.1.1 Caso 1 – Heurísticas do Grupo A

A Tabela 16 apresenta um resumo dos resultados obtidos para a aplicação individual de heurísticas do Grupo A. Como podemos observar, a média geral dos *scores* foi alta (5,62). Verificamos também que as heurísticas com o pior desempenho são as do subgrupo HA.7, i.e., aquelas que contemplam o RL ‘mod’, que geralmente corresponde a formações do tipo conceito-modificador, ou substantivo-adjetivo, no português. Entretanto, este desempenho é muito bom, especialmente considerando-se que, para o Corpus UNU, os resultados deste subgrupo foram mais baixos. Este resultado pode ser devido ao nosso esforço de trabalhar com códigos UNL consistentes: o Corpus Theses foi codificado por nós mesmos, após o esclarecimento de algumas definições da UNL, enquanto o Corpus UNU foi derivado da especificação original por diversos times UNL, introduzindo maior diversidade de interpretações e, logo, também inconsistências no uso dos RLs.

Analisando cada métrica separadamente (Tabela 17), verificamos que 81,77% dos sumários preservaram totalmente o foco do sumário, enquanto 78,82% dos sumários foram considerados bem formados. Verificamos ainda que cada heurística do Grupo A comprime, em média, cerca de 7,71% da sentença-fonte.

Por fim, verificamos que as heurísticas de alguns subgrupos não foram aplicadas nenhuma vez, por exemplo, as heurísticas que contemplam o RL ‘met’, que corresponde à relação semântica *method*, ou método necessário para a realização de uma ação ou evento. Analisando manualmente o corpus de textos, verificamos que apenas uma sentença contém o RL ‘met’ e, nesse único caso, a informação não era independente, ou seja, só poderia ser excluída por uma heurística do Grupo B. Os subgrupos HA.1 (RL ‘met’) e HA.8 (RL ‘nam’, que contempla a relação semântica *name*) tiveram aplicabilidade 0 neste caso.

Subgrupo	RL relacionado	Número de Heurísticas aplicadas	Média de compressão	Média de avaliação
HA.2	man	7	9,21%	5,75
HA.3	plc	9	8%	6
HA.4	aoj	8	5,88%	5,5
HA.5	tim	12	6,06%	5,95
HA.6	pos	7	8,91%	5,42
HA.7	mod	109	7,84%	5,11
HA.9	src	4	7,08%	5,5
HA.10	qua	8	9,38%	5,63
HA.11	cnt	6	7,01%	5,75
Totais	9	170		
Médias			7,71%	5,62

Tabela 16: Estatísticas de poda para o Caso 1

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	103	22	9	78,82%
Prejudicada	36	0	0	21,18%
Porcentagem	81,77%	12,94%	5,29%	

Tabela 17: Resultados para preservação do foco e textualidade no Caso 1

Exemplo de Aplicação de Heurísticas do Grupo A

A sentença [9] “*A ferramenta DB-LiOSE (DataBase - Link Oriented System For Education), implementada neste trabalho, possibilita que automaticamente se verifique os indicativos de qualidade definidos.*”, cuja codificação UNL é apresentada no Exemplo 16, foi sumarizada pela heurística HA.2.2, cujo subgrupo contempla a relação semântica *manner*, rotulada pelo RL ‘man’. O sumário UNL resultante é apresentado no Exemplo 17 e sua decodificação automática leva à sentença “*A ferramenta DB-LiOSE (DataBase - Link Oriented System For Education), implementada neste trabalho, possibilita que se verifique os indicativos de qualidade definidos.*”. Como podemos observar, a sentença continua apresentando elos coesivos entre os componentes sentenciais, fazendo com que tal sumário seja bem formado. Além disso, a informação principal da sentença-fonte, que é a função da ferramenta DB-LiOSE, também é preservada e, portanto, o sumário foi avaliado com o *score* máximo 6.

```
[S]
agt(allow.@entry,tool.@def)
mod(tool.@def,DB-LiOSE)
cnt(DB-LiOSE,DataBase - Link Oriented
  System For Education)
obj(implement.@past,tool.@def)
plc(implement.@past,work)
mod(work,this)
gol(allow.@entry,verify)
man(verify,automatically)
obj(verify,indicative.@pl.@def)
mod(indicative.@pl.@def,quality)
obj(define,indicative.@pl.@def)
[/S]
```

Exemplo 16: Codificação UNL de [9]

```
[S]
agt(allow.@entry,tool.@def)
mod(tool.@def,DB-LiOSE)
cnt(DB-LiOSE,DataBase - Link Oriented
  System For Education)
obj(implement.@past,tool.@def)
plc(implement.@past,work)
mod(work,this)
gol(allow.@entry,verify)
obj(verify,indicative.@pl.@def)
mod(indicative.@pl.@def,quality)
obj(define,indicative.@pl.@def)
[/S]
```

Exemplo 17: Possível sumário UNL de [9]

4.2.1.2 Caso 2 – Heurísticas do Grupo B

A Tabela 18 resume os resultados para a aplicação individual de heurísticas do Grupo B. Considerando as médias gerais dos *scores* obtidos, todos os subgrupos de heurísticas são aceitáveis, com uma média geral de 5,01. Analisando cada métrica separadamente (Tabela 19), verificamos que 68,52% dos sumários preservaram totalmente o foco do sumário, enquanto 75,93% foram considerados bem formados. Cada heurística do Grupo B comprime, em média, cerca de 32,42% da sentença-fonte.

Verificamos também que as heurísticas de três subgrupos não foram aplicadas nenhuma vez.: HB.16 (RL ‘con’, que contempla a relação semântica *condition*), HB.17 (RL ‘agt’, que contempla a relação semântica *agent*) e HB.18 (RL ‘ben’, que contempla a relação semântica *beneficiary*). Para investigar o porquê da não aplicabilidade desses subgrupos, analisamos manualmente o corpus de textos. Em relação ao subgrupo HB.16, verificamos que não houve nenhuma ocorrência do RL ‘con’ em todo o corpus. Já em relação ao subgrupo HB.17 verificamos que, embora o RL ‘agt’ tenha ocorrências no corpus, nenhuma das ocorrências satisfaz as condições contextuais das heurísticas. Esse resultado é apropriado, pois a relação de agente é essencial na maioria dos casos e, portanto, as heurísticas que contemplam essa relação foram definidas de forma a serem aplicáveis em alguns poucos casos, nos quais asseguradamente as RBs representam informações não essenciais¹⁰. Por fim, em relação ao subgrupo HB.18 verificamos que, embora o RL ‘ben’ tenha ocorrências no corpus, na maioria dos casos ele representa informações independentes e, portanto, estas deveriam ser contempladas por uma heurística do Grupo A.

¹⁰ Exemplos da aplicação correta de heurísticas do subgrupo HB.17 podem ser encontradas em (Martins e Rino,2002).

Subgrupo	RL relacionado	Número de Heurísticas aplicadas	Média de compressão	Média de avaliação
HB.1	met	1	23,08%	4
HB.2	man	14	34,99%	5,25
HB.3	plc	18	19,95%	5,35
HB.4	aoj	15	33,23%	3,75
HB.5	tim	1	50%	6
HB.7	mod	31	21,86%	4,39
HB.9	src	12	14,12%	5,5
HB.12	scn	3	15,38%	5,25
HB.13	rsn	5	61,66%	6
HB.14	pur	6	52,39%	5,6
HB.15	dur	1	30%	4
Totais	11	107		
Médias			32,42%	5,01

Tabela 18: Estatísticas de poda para o Caso 2

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	63	13	6	75,93%
Prejudicada	11	9	6	24,07%
Porcentagem	68,52%	20,37%	11,11%	

Tabela 19: Resultados para preservação do foco e textualidade no Caso 2

Exemplo de Aplicação de Heurísticas do Grupo B

A sentença [10] “*Segundo Hatzmanikatis et al. (1995), a avaliação de um hiperdocumento pode surgir das práticas da área de Engenharia de Software e de Métricas de Software.*”, cuja codificação UNL é apresentada no Exemplo 18, foi resumida pela heurística HB.7.3, que contempla a relação entre conceito e modificador, representada em UNL pelo RL ‘mod’. O sumário UNL correspondente é apresentado no Exemplo 19 e sua decodificação automática leva à sentença “*Segundo Hatzmanikatis et al. (1995), a avaliação de um hiperdocumento pode surgir das práticas.*”. Pode-se considerar que, neste exemplo, a informação principal não foi mantida, apesar de o sumário continuar sendo uma sentença bem formada. Assim, este sumário foi avaliado com o *score* 2.

```
[S]
agt(state.@past,Hatzmanikatis)
tim(state.@past,1995.@parenthesis)
obj(state.@past,:01)
mod:01(evaluation.@entry,
hyperdocument.@indef)
aoj(arise.@may,evaluation.@entry)
src(arise.@may,practice.@pl)
mod(practice.@pl,:02)
and:02(Software Engineering,Software
Measure.@pl)
[/S]
```

Exemplo 18: Codificação UNL de [10]

```
[S]
agt(state.@past,Hatzmanikatis)
tim(state.@past,1995.@parenthesis)
obj(state.@past,:01)
mod:01(evaluation.@entry,
hyperdocument.@indef)
aoj(arise.@may,evaluation.@entry)
src(arise.@may,practice.@pl)
[/S]
```

Exemplo 19: Possível sumário UNL de [10]

4.2.1.3 Crítica Parcial

A análise parcial dos casos 1 e 2 leva às seguintes observações em relação às hipóteses 1 e 2:

- 1) *Aplicabilidade* - A maioria das heurísticas de ambos os grupos é aplicável. Houve apenas dois subgrupos de heurísticas do Grupo A e três subgrupos de heurísticas do Grupo B com aplicabilidade 0, devido às características do corpus utilizado.
- 2) *Satisfatibilidade* individual - Todas as heurísticas aplicadas apresentaram resultados satisfatórios, i.e., cada heurística produziu pelo menos um sumário satisfatório. Além disso, todos os subgrupos de heurísticas obtiveram uma média dos *scores* muito acima de média - 3,5. Verificamos também que a satisfatibilidade das heurísticas do Grupo A é maior do que as do Grupo B.

4.2.2 Aplicação conjunta de heurísticas

4.2.2.1 Caso 3 – Heurísticas do Grupo A

A Tabela 20 mostra alguns exemplos de resultados para a aplicação conjunta de heurísticas do Grupo A a 51 sentenças do Corpus Theses. De forma semelhante à adotada na compilação dos resultados para o corpus UNU, as heurísticas são apresentadas nessa tabela por sua ordem de aplicação. Avaliando-se os resultados da aplicação conjunta de heurísticas do Grupo A a todas as sentenças do Corpus Theses, obtivemos uma média dos *scores* de 4,71. Analisando cada métrica separadamente (Tabela 21), verificamos que a maioria dos sumários preservou o foco (60,78%) e resultou em uma sentença bem formada (64,71%), para uma taxa de compressão média de 27,39%.

Sentença	Heurísticas aplicadas	Compressão	Score
Daniela s3	HA.10.1, HA.5.6, HA.10.1	37,5%	6
Daniela s4	HA.7.6	7%	6
Daniela s5	HA.7.7	12%	6
Daniela s6	HA.5.4, HA.7.1, HA.7.1, HA.5.6	50%	4
Daniela s7	HA.5.3, HA.6.1, HA.7.6, HA.7.1, HA.3.3, HA.7.1, HA.7.1	33,33%	6
Daniela s8	HA.7.5, HA.3.3	16,67%	6
Daniela s9	HA.7.12	8%	5
...
Totais	170 heurísticas		
Médias		27,39%	4,71

Tabela 20: Estatísticas de poda parciais para o Caso 3

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	20	10	3	64,71%
Prejudicada	11	6	1	35,29%
Porcentagem	60,78%	31,37%	7,84%	

Tabela 21: Resultados para preservação do foco e textualidade no Caso 3

Exemplo da Aplicação Conjunta de Heurísticas do Grupo A

A sentença [11] “*Dentre as diversas áreas que estão em expansão atualmente, podemos destacar duas delas: a Internet e a de Qualidade.*”, referenciada por “Daniela s3” na Tabela 20, foi sumarizada por três heurísticas (vide coluna *Heurísticas Aplicadas*). Sua codificação UNL e correspondente sumário são apresentados nos Exemplos 20 e 21, respectivamente. A decodificação automática desse sumário UNL leva à sentença “*Dentre as áreas que estão em expansão, podemos destacar a de Internet e a de Qualidade.*”. Esta sentença é bem formada, pois apresenta elos coesivos e é coerente, permitindo, assim, a veiculação de um significado. Além disso, ela exclui somente informações periféricas ou que podem ser inferidas pelo leitor e, portanto, preserva o foco da sentença-fonte. Logo, o sumário foi avaliado com o *score* máximo 6.

```
[S]
qua(area.@pl.@def,various)
agt(area.@pl.@def,
  expand.@progress.@entry)
tim(expand@progress.@entry,nowadays)
pof(:01,area@pl.@def)
and:01(Internet,quality)
agt(we,highlight)
obj(highlight,:01)
qua(highlight,two)
[/S]
```

Exemplo 20: Codificação UNL de [11]

```
[S]
agt(area.@pl.@def,
  expand.@progress.@entry)
pof(:01,area@pl.@def)
and:01(Internet,quality)
agt(we,highlight)
obj(highlight,:01)
[/S]
```

Exemplo 21: Possível sumário UNL de [11]

4.2.2.2 Caso 4 – Heurísticas do Grupo B

A Tabela 22 resume os resultados para a aplicação conjunta de heurísticas do Grupo B a 43 sentenças do Corpus Theses, totalizando uma média dos *scores* de 4,58. Apesar desse bom resultado, observamos que, na maioria dos casos em que a avaliação foi abaixo de 3,5, as podas se deveram a heurísticas dos subgrupos HB.4 e HB.7, correspondentes aos RLs ‘aoj’ e ‘mod’, respectivamente. Analisando cada métrica separadamente, verificamos que a maioria dos sumários preservou totalmente o foco da sentença-fonte (54,76%) e foi considerada bem formada, apresentando um índice de textualidade de 69,05% (Tabela 23). Verificamos, ainda, que há uma relação entre a taxa de compressão e o índice de avaliação dos sumários: quanto maior essa taxa, menor o *score* da sentença correspondente. A média geral das taxas de compressão das sentenças deste corpus foi de 49,19%.

Sentença	Heurísticas aplicadas	Compressão	Score
Daniela s4	HB.7.3, HB.9.3	58,33%	3
Daniela s5	HB.3.5	37,5%	6
Daniela s7	HB.3.5, HB.13.3	76,19%	6
Daniela s8	HB.7.4, HB.4.3, HB.2.3	50%	1
Daniela s9	HB.4.1	16,67%	5
Daniela s10	HB.4.3	44,44%	1
Daniela s11	HB.4.3	50%	1
...
Totais	77 heurísticas, 43 sentenças		
Médias		49,19%	4,58

Tabela 22: Estatísticas de poda parciais para o Caso 4

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	19	9	1	69,05%
Prejudicada	4	5	4	30,95%
Porcentagem	54,76%	33,33%	11,91%	

Tabela 23: Resultados para preservação do foco e textualidade no Caso 4

Exemplo da Aplicação Conjunta de Heurísticas do Grupo B

A sentença [12] “*Observa-se que da mesma maneira que o número de usuários da Internet aumenta, os hiperdocumentos que estão disponibilizados nela também crescem em quantidade e conteúdo.*”, referenciada na Tabela 22 por *Daniela s8*, cuja codificação UNL é ilustrada no Exemplo 22, foi sumarizada por 3 heurísticas, correspondentes às relações semânticas *modifier* (RL ‘mod’), *attribute of object* (RL ‘aoj’) e *manner* (RL ‘man’). O sumário UNL resultante é apresentado no Exemplo 23 e sua decodificação automática levou à sentença “*Observa-se que da mesma forma que o número aumenta, os hiperdocumentos crescem*”. Claramente, a textualidade do sumário apresentado foi prejudicada, devido à incompletude do sintagma “*número de usuários*” no sumário. Além disso, podemos considerar que a informação principal não foi preservada, pois o sumário omite o tipo de crescimento sofrido pelos hiperdocumentos. Sumários desse tipo são totalmente inadequados e foram avaliados com o *score* 1.

```
[S]
obj(observed,increase.@entry)
agt(increase.@entry,number.@def)
mod(number.@def,user.@pl)
mod(user.@pl,Internet)
man(increase.@entry,way)
aoj(same,way)
bas(same,grow)
agt(grow,hyperdocument.@pl.@def)
aoj(dispose,hyperdocument.@pl.@def)
plc(dispose,Internet.@def)
man(grow,:01)
and:01(quantity,content)
[/S]
```

Exemplo 22: Codificação UNL de [12]

```
[S]
obj(observed,increase.@entry)
agt(increase.@entry,number.@def)
man(increase.@entry,way)
aoj(same,way)
bas(same,grow)
agt(grow,hyperdocument.@pl.@def)
[/S]
```

Exemplo 23: Possível sumário UNL de [12]

4.2.2.3 Crítica Parcial da Aplicação Conjunta de Heurísticas dos Grupos A e B

Os casos 3 e 4 referem-se às hipóteses 1, 3 e 4 descritas na seção 3.2. Assim, temos o seguinte:

- 1) *Aplicabilidade* - Para este corpus, as heurísticas mantêm sua aplicabilidade quando aplicadas individualmente ou em conjunto a uma sentença.
- 2) *Satisfatibilidade conjunta* - A aplicação conjunta tanto de heurísticas do Grupo A quanto do Grupo B é satisfatória. Verificamos também que, conforme observado com o Corpus UNU, a satisfatibilidade é maior aplicando-se heurísticas do Grupo A, o que sugere que, ao se optar por apenas um grupo de heurísticas, o Grupo A deve ser priorizado.
- 3) *Consistência* - A aplicação conjunta de heurísticas foi consistente para todos os dados de entrada.

Por fim, uma observação não prevista por nossas hipóteses está relacionada às taxas de compressão: *scores* mais baixos tendem a estar ligados a altas taxas de compressão.

4.2.3 Ordem de Aplicação

4.2.3.1 Caso 5 – Aplicar primeiro heurísticas do Grupo A

A Tabela 24 resume os valores obtidos ao se privilegiar a poda por heurísticas do Grupo A, seguidas da poda por heurísticas do Grupo B, mostrando os valores de algumas das 51 sentenças sumarizadas neste caso. Avaliando-se os resultados gerais, obtivemos uma média dos *scores* de 4,04 e uma média das taxas de compressão de 42,56%. Analisando cada métrica separadamente (Tabela 25), verificamos que, apesar da média dos *scores* ter sido aceitável, apenas 39,22% dos sumários preservaram totalmente o foco da sentença-fonte. Por outro lado, 60,78% dos sumários resultaram em sentenças bem formadas.

Sentença	Grupo A	Grupo B	Total	Compressão	Score
Daniela s3	3	0	3	37,5%	6
Daniela s4	1	2	3	66,67%	2
Daniela s5	1	1	2	50%	4
Daniela s6	4	0	4	50%	4
Daniela s7	7	1	8	42,86%	6
Daniela s8	2	1	3	66,67%	1
Daniela s9	1	1	2	25%	3
...
Totais	172	35	207		
Médias				42,56%	4,04

Tabela 24: Estatísticas de poda parciais para o Caso 5

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	13	13	5	60,78%
Prejudicada	7	9	4	39,22%
Porcentagem	39,22%	43,14%	17,65%	

Tabela 25: Resultados para preservação do foco e textualidade no Caso 5

Exemplo da Aplicação de Heurísticas do Grupo A, depois do Grupo B

A sentença [13] “*A Internet no princípio era utilizada por militares norte-americanos e posteriormente por acadêmicos e usuários privilegiados.*”, referenciada na Tabela 24 por Daniela s6 e cuja codificação UNL é apresentada no Exemplo 24, foi resumida por 4 heurísticas. O sumário UNL resultante é apresentado no Exemplo 25 e sua decodificação automática leva à sentença “*A Internet era utilizada por militares, acadêmicos e usuários.*”. Analisando a sentença-fonte, verificamos que a sentença resumida pode ser considerada bem formada, mas não preserva totalmente sua idéia principal, por não deixar claro o tempo em que essas informações são verdadeiras (*no princípio e posteriormente*). Sumários desse tipo foram avaliados com o *score* 4.

```
[S]
obj(utilize.@past,Internet.@entry)
agt(utilize.@past,:01)
and:01(militar.@pl,:02)
tim(militar.@pl,beginning)
mod(militar.@pl,North-American)
and:02(academic.@pl,user.@pl)
mod(user.@pl,privileged.@pl)
tim(:02,later)
[/S]
```

Exemplo 24: Codificação UNL de [13]

```
[S]
obj(utilize.@past,Internet.@entry)
agt(utilize.@past,:01)
and:01(militar.@pl,:02)
and:02(academic.@pl,user.@pl)
[/S]
```

Exemplo 25: Possível sumário UNL de [13]

4.2.3.2 Caso 6 – Aplicar primeiro heurísticas do Grupo B

A Tabela 26 resume os valores obtidos ao se privilegiar a poda por heurísticas do Grupo B, seguidas da poda por heurísticas do Grupo A. Verificamos que a média geral foi aceitável (3,76), embora baixa, enquanto a média de compressão resultou bastante alta em relação aos outros casos (54,13%). Observamos também que, de um modo geral, os piores resultados correspondem às maiores taxas de compressão. Verificamos, ainda, que a maioria dos sumários é bem formada (índice de textualidade de 68,63%), mas apenas 33,33% dos sumários preservam totalmente o foco da sentença-fonte (Tabela 27).

Sentença	Grupo A	Grupo B	Total	Compressão	Score
Daniela s3	3	0	3	37,5%	6
Daniela s4	1	2	3	66,67%	2
Daniela s5	1	1	2	50%	4
Daniela s6	4	0	4	50%	4
Daniela s7	1	2	3	80,95%	6
Daniela s8	0	3	3	50%	2
Daniela s9	1	1	2	25%	3
...
Totais – 51	68	77	145		
Médias				54,13%	3,76

Tabela 26: Estatísticas de poda parciais para o Caso 6

Foco	Preservado	Preservado Parcialmente	Ausente	Porcentagem
Textualidade				
Preservada	13	14	8	68,63%
Prejudicada	4	5	7	31,37%
Porcentagem	33,33%	37,25%	29,41%	

Tabela 27: Resultados para preservação do foco e textualidade no Caso 6

Exemplo da Aplicação de Heurísticas do Grupo B, depois do Grupo A

A sentença [14] “*A área educacional tenta inovar o ensino tradicional e busca essas inovações através da utilização de novas tecnologias (Campos et al., 1999).*”, cuja codificação UNL é apresentada no Exemplo 26, Ela foi sumarizada pelas heurísticas HB.1.1, HB.9.3, HA.7.12 e HA.7.7, correspondentes às relações semânticas *method* (RL ‘met’), *source* (RL ‘src’) e *modifier* (RL ‘mod’). O sumário UNL resultante é apresentado no Exemplo 27 e sua decodificação automática leva à sentença “*A área tenta inovar o ensino e busca inovações*”. Embora a sentença seja bem formada, não preserva o foco da sentença fonte, pois, ao excluir *educacional*, o sumário omite parte essencial do conceito correspondente ao sujeito da sentença. Essa sentença, portanto, foi avaliada com o *score* 2. Podemos verificar também, no Exemplo 26, que a informação correspondente a *educacional* é representada pelo RL ‘mod’ e foi excluída pela heurística HA.7.12.

```
[S]
and:03(try.@entry,search)
agt:01(:03,area.@def)
mod:01(area.@def,educational)
gol:01(try.@entry,innovate)
obj:01(innovate,teaching.@def)
mod:01(teaching.@def,traditional)
obj:01(search,innovation.@pl)
met:01(search,utilization)
obj:01(utilization,technology.@pl)
mod:01(technology.@pl,new)
src:01:,02.@parenthesis)
tim:02(Campos et al.,1999)
[/S]
```

Exemplo 26: Codificação UNL de [14]

```
[S]
and:03(try.@entry,search)
agt:01(:03,area.@def)
gol:01(try.@entry,innovate)
obj:01(innovate,teaching.@def)
obj:01(search,innovation.@pl)
[/S]
```

Exemplo 27: Possível sumário UNL de [14]

4.2.3.3 Crítica Parcial sobre a Ordem de Aplicação de Heurísticas dos Grupos A e B

Os casos 5 e 6 podem ser analisados em relação às hipóteses 1, 4 e 5 descritas na seção 3.2. Assim, temos o seguinte:

- 1) *Aplicabilidade* - Verificamos que as heurísticas são aplicáveis independentemente da ordem de aplicação.
- 2) *Consistência* - A aplicação conjunta de heurísticas foi consistente para todos os dados de entrada, para qualquer ordem de aplicação.
- 3) *Interdependência* - Verificamos que realmente há uma interdependência entre as heurísticas do Grupo A e do Grupo B, pois os resultados obtidos nos casos 5 e 6 foram significativamente diferentes. Os dados indicam que, devido a essa interdependência, a melhor estratégia para poda é aplicar primeiro todas as heurísticas do Grupo A possíveis e, depois, as do Grupo B.

4.2.4 Crítica Geral Sobre o Corpus Theses

O estudo de casos realizado para o Corpus Theses (resumido na Tabela 28) mostrou resultados similares aos obtidos com o Corpus UNU, tanto em relação às hipóteses quanto em relação à taxa de compressão e sua relação com as avaliações gerais.

	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6
Média dos Scores	5,62	5,01	4,71	4,58	4,04	3,76
Compressão	7,71%	32,42%	27,39%	49,19%	42,56%	54,13%

Tabela 28: Estatísticas de poda global para o Corpus Theses

1) Aplicabilidade – A maioria das heurísticas de ambos os grupos foi considerada aplicável. No entanto, os subgrupos de heurísticas HA.1 (RL ‘met), HA.8 (RL ‘nam’), HB.16 (RL ‘con’), HB.17 (RL ‘agt’) e HB.18 (RL ‘ben’) obtiveram aplicabilidade 0 neste corpus.

2) Satisfatibilidade individual - Todos os subgrupos de heurísticas apresentam uma média dos *scores* satisfatória, i.e., acima de 3,5. No entanto, verificamos também que a maioria dos sumários com *score* abaixo de 3,5 foi produzida por heurísticas que contemplam os RLS ‘aoj’ (como nos Exemplos 22 e 23) e ‘mod’ (como nos Exemplos 18 e 19).

3) Satisfatibilidade conjunta – A aplicação conjunta de heurísticas é satisfatória, pois em todos os casos (casos de 3 a 6), a maioria dos sumários foi considerada satisfatória, i.e., foi avaliada com um *score* acima de 3,5. Verificamos também que a satisfatibilidade é maior aplicando-se heurísticas do Grupo A. Além disso, em todos os casos os resultados para textualidade foram melhores do que os resultados para a preservação do foco. Por exemplo, a Tabela 21 mostra que, embora 64,71% dos sumários produzidos pela aplicação conjunta de heurísticas do Grupo A preservem a textualidade, apenas 60,78% preservam totalmente o foco. Já a Tabela 23 mostra que, embora 69,05% dos sumários produzidos pela aplicação conjunta de heurísticas do Grupo B preservem a textualidade, apenas 54,76% preservam totalmente o foco. Como consideramos a preservação da idéia central mais importante do que a preservação da textualidade, as heurísticas do Grupo A tiveram um desempenho melhor. Assim, ao se optar por apenas um grupo de heurísticas, o Grupo A deve ser priorizado, pois tende a garantir sumários mais satisfatórios, principalmente em relação à preservação do foco.

4) Consistência - A aplicação conjunta de heurísticas provou ser consistente em todos os casos, independentemente do grupo de heurísticas ou de sua ordem de aplicação.

5) Interdependência – Como já discutimos anteriormente, a ordem de aplicação de heurísticas de cada grupo é importante, pois, dependendo da ordem adotada, diferentes heurísticas são aplicadas e, conseqüentemente, o resultado pode variar bastante. Nossos resultados sugerem que, para se obter melhores sumários, o Grupo A deve ser privilegiado.

Por fim, observamos que existe uma dependência entre a taxa de compressão e a qualidade do sumário. De forma similar à observada para o Corpus UNU, quanto maior a taxa de compressão, pior o resultado (Gráfico 2). No entanto, verificamos que, para taxas de compressão a partir de 71%, os resultados obtidos para este corpus foram surpreendentemente altos.

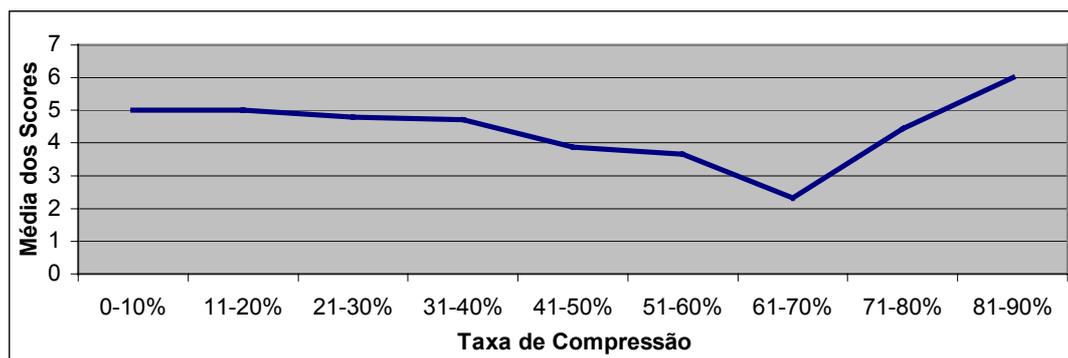


Gráfico 2: Desempenho dos sumários em relação à taxa de compressão

Analisando separadamente esses sumários, observamos primeiramente que houve poucos deles com tal taxa de compressão (apenas 21), o que pode ter levado a esse resultado diferente das observações anteriores, ou seja, nestes sumários altas taxas de compressão produziram bons sumários. Verificamos também que tais sumários só ocorreram nos casos 2, 4 e 6, ou seja, quando foram aplicadas apenas heurísticas do Grupo B (casos 2 e 4) ou quando as heurísticas do Grupo B tiveram prioridade (caso 6). Além disso, verificamos que, para taxas de compressão acima de 70%, de um modo geral, sumários com *scores* altos resultaram da aplicação de poucas heurísticas, enquanto sumários com *scores* baixos foram produzidos por um grande número de heurísticas. A Tabela 29 apresenta o número médio de heurísticas aplicadas a sumários avaliados com cada *score*. Como podemos observar, sumários com *score* 6 foram produzidos, em média, por apenas 1,55 heurísticas. Já os sumários com qualquer outro *score* foram produzidos por um número médio de heurísticas bem mais alto. Esses dados sugerem que, para as heurísticas do Grupo B, para gerar melhores resultados é importante controlar o número de heurísticas aplicadas, sendo pouco relevante a taxa de compressão alcançada.

	<i>Score</i> 6	<i>Score</i> 5	<i>Score</i> 4	<i>Score</i> 3	<i>Score</i> 1
Número médio de heurísticas aplicadas por sentença	1,55	5	4	3	4,33

Tabela 29: Estatísticas para sentenças com taxas de compressão acima de 70%

Outra possível explicação para esse resultado está no fato de este corpus apresentar sentenças mais complexas do que as do Corpus UNU e, por essa razão, haver maior probabilidade de se identificar e excluir mais informações pouco relevantes, levando a bons sumários, mesmo que altamente condensados. No entanto, devido à baixa ocorrência desses casos, ambas as hipóteses deverão ser investigadas futuramente.

5 Conclusões Gerais do Estudo de Casos

Analisando os resultados do estudo de casos para ambos os corpora, concluímos que:

- 1) A maioria das heurísticas foi considerada aplicável aos corpora considerados. No entanto, os subgrupos HA.1 (RL ‘met’), HA.8 (RL ‘nam’), HB.16 (RL ‘con’), HB.17 (RL ‘agt’) e HB.18 (RL ‘ben’) não foram aplicáveis a nenhum dos dois corpora. Embora esses resultados pareçam indicar que esses subgrupos de heurísticas não possuem utilidade para a sumarização, devemos lembrar que os corpora analisados são limitados a domínios e gêneros específicos e, portanto, os resultados não podem ser considerados conclusivos. Assim, estudos posteriores deverão ser realizados com diferentes corpora para, então, decidir se elas devem ou não ser mantidas no elenco de heurísticas do UNLSumm.
- 2) A maioria dos subgrupos de heurísticas levou a sumários satisfatórios. No entanto, alguns deles levaram a muitos resultados insatisfatórios. Para o Corpus UNU, os subgrupos de heurísticas responsáveis pelos baixos desempenhos são aqueles que contemplam os RLs ‘aoj’ (HA.4 e HB.4), ‘mod’ (HA.7 e HB.7) e ‘pur’ (HB.14). Já para o Corpus Theses, os que levaram aos baixos desempenhos foram aqueles relacionados aos RLs ‘aoj’, ‘mod’ e ‘src’ (HA.9 e HB.9). Portanto, os RLs mais problemáticos nesse estudo foram ‘aoj’ e ‘mod’. De fato, as relações conceituais de ‘atributo de objeto’ e ‘modificador’, relativas a esses RLs, podem ser consideradas essenciais no português, como observou um lingüista especialista em UNL. Evidências disso foram dadas anteriormente, onde sumários produzidos pela mesma heurística possuem qualidade bastante diversa. Por exemplo, a sentença UNL correspondente à sentença [15] apresentada no Exemplo 28¹¹ foi sumarizada no Caso 1 pela heurística HA.7.5, que contempla o RL ‘mod’, gerando um sumário UNL cuja decodificação automática é apresentada no Exemplo 29, tendo sido avaliada com o *score* 6. A heurística HA.7.5, neste caso, excluiu a palavra “current”, pouco proeminente nesta sentença.

“The Conference of Directors (CONDIR) congregates the Directors of the RTCs and RTPs and it is called by to review and evaluate programmes of research being undertaken, and assist the Rector in the improvement of **current** programmes.”

Exemplo 28: Sentença [15]

“The Conference of Directors (CONDIR) congregates the Directors of the RTCs and RTPs and it is called by to review and evaluate programmes of research being undertaken, and assist the Rector in the improvement of programmes.”

Exemplo 29: Possível sumário de [15]

Já a sentença UNL correspondente à sentença [16], apresentada no Exemplo 30, apesar de ter sido podada no Caso 1 também pela heurística HA.7.5, gerou um sumário UNL que foi avaliado com o *score* 2 (o Exemplo 31 mostra sua decodificação automática).

¹¹ Os exemplos 2, 3, 4 e 5 foram extraídos dos estudos de casos realizados para o Corpus UNU.

Neste caso, vemos que a exclusão da RB rotulada por ‘mod’, que representava o conceito “nuclear”, modificou o sentido da sentença original.

“The IAEA promotes peaceful uses of **nuclear** energy by serving as a intergovernmental forum for scientific and technical cooperation.”

Exemplo 30: Sentença [16]

“The IAEA promotes peaceful uses of energy by serving as a intergovernmental forum for scientific and technical cooperation.”

Exemplo 31: Possível sumário de [16]

Ou seja, este exemplo mostra que, embora na sentença [15] a informação representada por ‘mod’ seja pouco relevante, na sentença [16] a informação é essencial. Assim, não podemos garantir que as heurísticas relacionadas a ‘mod’ excluirão sempre informações pouco relevantes, o mesmo valendo para o RL ‘aoj’. Ainda em relação a esses dois subgrupos de heurísticas, verificamos que sua aplicação conjunta também prejudica a qualidade dos sumários. Isso ficou evidente ao observarmos que, na aplicação conjunta de heurísticas, a maioria dos sumários insatisfatórios foi gerada por pelo menos uma dessas heurísticas. Dadas essas evidências, as heurísticas envolvendo os RLs ‘aoj’ e ‘mod’ deverão ser excluídas do elenco de poda do UNLSumm.

- 3) As heurísticas produzem resultados satisfatórios quando aplicadas conjuntamente (casos 3 a 6). No entanto, verificamos que os *scores* são mais baixos e as taxas de compressão mais altas do que quando se aplica unicamente uma heurística a uma sentença, como podemos verificar nas Tabelas 15 e 28, reproduzidas abaixo. Isso indica uma possível relação entre taxas de compressão altas e *scores* baixos, como discutiremos adiante.

	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6
Média dos Scores	5,6	4,43	4,3	3,58	3,58	2,85
Compressão	7,71%	37,23%	26,36%	48,03%	40,41%	56,06%

Tabela 15: Estatística de poda global para o Corpus UNU

	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6
Média dos Scores	5,62	5,01	4,71	4,58	4,04	3,76
Compressão	7,71%	32,42%	27,39%	49,19%	42,56%	54,13%

Tabela 28: Estatísticas de poda global para o Corpus Theses

- 4) O sistema é totalmente consistente. O UNLSumm conseguiu aplicar com sucesso as heurísticas a todas as sentenças em foco no estudo de casos, i.e., a todos os dados de ambos os corpora, para qualquer grupo de heurísticas ou ordem de aplicação, sem levar a conflitos na tomada de decisão.
- 5) Os grupos A e B de heurística apresentam uma interdependência de aplicação, sendo que ao aplicarmos primeiro todas as heurísticas do Grupo A (como realizado no Caso 5), são produzidos sumários de maior qualidade. Esse resultado foi verificado em ambos os corpora e, portanto, ele sugere que, quando ambos os grupos de heurísticas forem aplicados, o UNLSumm deverá priorizar o Grupo A.

- 6) Verificamos que a qualidade dos sumários cai à medida que a taxa de compressão aumenta. Observando os resultados obtidos com ambos os corpora (Gráfico 3), verificamos que, a partir de cerca de 41% de taxa de compressão para o Corpus UNU e 51% para o Corpus Theses os sumários tendem a apresentar resultados insatisfatórios. No entanto, após 70% de compressão, essas observações não se aplicam ao Corpus Theses. Por outro lado, como poucos sumários apresentaram taxa de compressão acima de 70%, não podemos chegar a uma conclusão clara sobre esses resultados, que deverão ser melhor investigados futuramente. No momento, tomando por base o fato de na maioria dos casos, os sumários com taxas de compressão acima de 45% serem insatisfatórios, limitaremos o UNLSumm à taxa de compressão máxima de 45%. Esse limite será novamente discutido no capítulo 7.

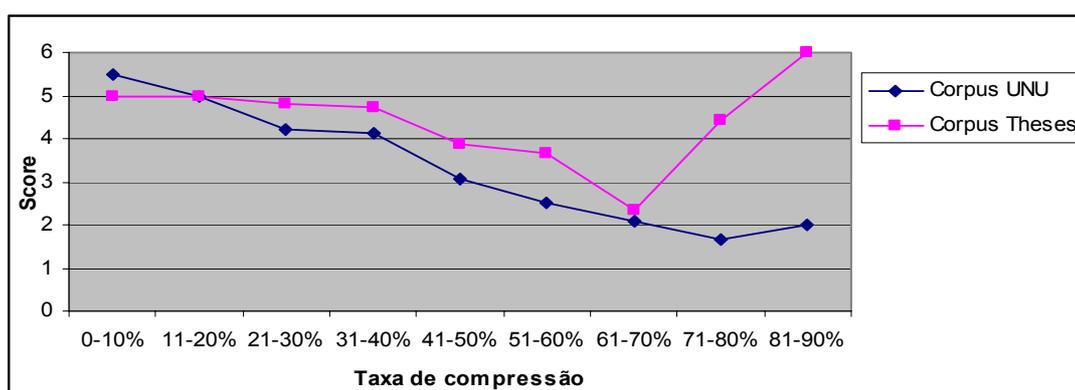


Gráfico 3: Taxa de compressão em relação ao *score*

- 7) Verificamos também, no estudo com ambos os corpora, que as heurísticas do Grupo A produzem sumários melhores do que as do Grupo B (como mostram as Tabelas 15 e 28, reproduzidas no item 3 desta seção). Os baixos *scores* dos sumários produzidos por heurísticas do Grupo B devem-se à baixa preservação do foco, embora correspondam a sentenças bem formadas, como podemos observar pelo Gráfico 4: somente em um dos casos – o Caso 2 para o Corpus Theses – houve avaliação próxima a 70%. Logo, ao se escolher a aplicação de apenas um dos grupos de heurísticas, deve-se privilegiar aquelas do Grupo A.

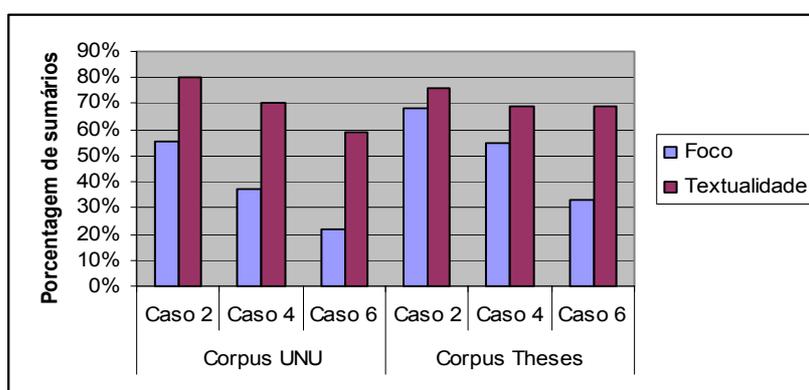


Gráfico 4: Preservação do foco e textualidade em sumários produzidos nos casos 2, 4 e 6

- 8) Por fim, em ambos os corpora, a aplicação conjunta de heurísticas de um só grupo (casos 3 e 4) levou a resultados melhores do que a aplicação conjunta de heurísticas de ambos os grupos (casos 5 e 6), como podemos observar nas Tabelas 15 e 28.
- 9) As considerações anteriores nos levam a uma conclusão final, indicando a melhor abordagem a ser utilizada pelo UNLSumm na sumarização de sentenças. Como vimos em 7), as heurísticas do Grupo A levam a melhores sumários, sugerindo que as melhores abordagens são aquelas adotadas nos casos 1, 3 e 5. Vimos também, em 4), que no caso 1 as taxas de compressão são muito baixas e, portanto, a abordagem utilizada nesse caso também não é a mais apropriada para a sumarização. Assim, as melhores abordagens seriam as adotadas nos casos 3 e 5. Analisando os resultados obtidos nesses casos, verificamos três fatores. Primeiro: por um lado, o caso 3 apresenta *scores* mais altos, como vimos em 8) e, por outro lado, apresenta taxas de compressão mais baixas, como vimos nas Tabelas 15 e 28. Segundo: no caso 5, apesar de os *scores* serem mais baixos, existe a vantagem de as taxas de compressão serem mais altas, mas com médias abaixo de 45% e, como vimos em 6), até esse valor em geral são gerados sumários satisfatórios. Terceiro: outra vantagem da abordagem utilizada no caso 5 é que também são aplicadas heurísticas do Grupo B, o que é desejável, pois assim pode-se cobrir mais casos possíveis de sumarização. Esses três fatores indicam, portanto, que, de forma a produzir sumários satisfatórios com uma taxa de compressão não muito baixa, o ideal é utilizar a abordagem utilizada no caso 5, isto é, aplicar primeiro heurísticas do Grupo A e depois do Grupo B. Portanto, essa abordagem deverá ser utilizada por *default*, no UNLSumm.

6 Readequação do Conjunto de Heurísticas

Considerando que os RLs ‘aoj’ e ‘mod’ são os mais problemáticos para gerar sumários de qualidade, os subgrupos que contemplam esses RLs foram excluídos do elenco de heurísticas, resultando no elenco versão 2.0 de heurísticas apresentado no Anexo 2. Para avaliar o desempenho desse novo conjunto de heurísticas, o estudo de casos foi refeito para os mesmos corpora. Comparando os resultados obtidos utilizando as versões 1.3 e 2.0, verificamos que, tanto para o Corpus UNU (Gráfico 5) quanto para o Corpus Theses (Gráfico 6), houve um acréscimo significativo nas médias dos *scores* obtidos nos casos 1-6. Esses dados indicam, ainda, que a versão 2.0 apresenta melhores resultados do que aqueles obtidos pela versão 1.3, descritos em detalhes neste relatório.

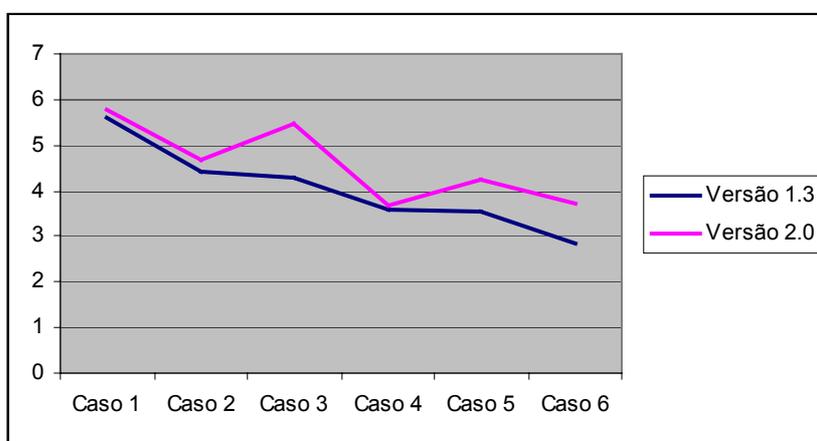


Gráfico 5: Comparação geral dos *scores* obtidos para o Corpus UNU

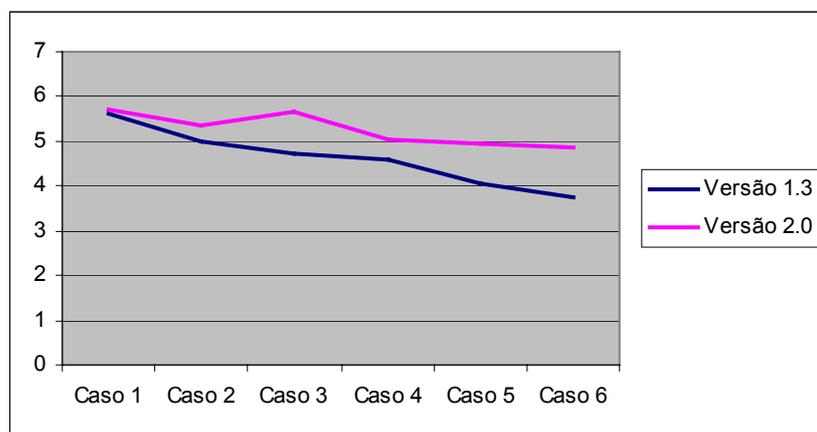


Gráfico 6: Comparação geral dos *scores* obtidos para o Corpus Theses

Analisando separadamente o foco e a textualidade, podemos observar que, na grande maioria dos casos, o percentual de sumários que mantêm o foco e preservam a textualidade aumentou, tanto para o Corpus UNU (Gráficos 7 e 8) quanto para o Corpus Theses (Gráficos 9 e 10). Para o Corpus UNU, os valores foram mais altos na maioria dos casos, tanto para a preservação do foco, quanto para a textualidade.

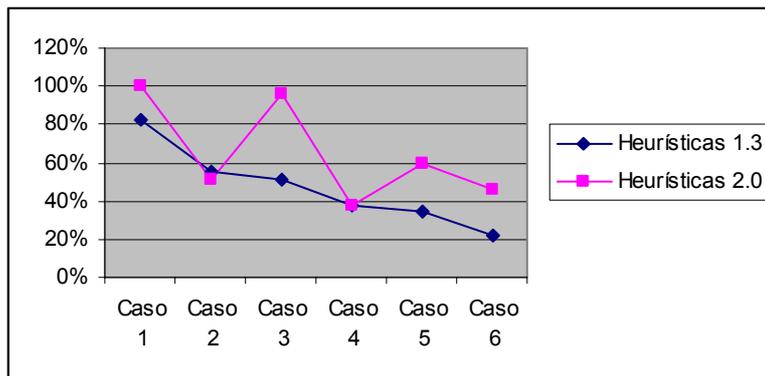


Gráfico 7: Percentual de sumários do Corpus UNU que preservam o foco

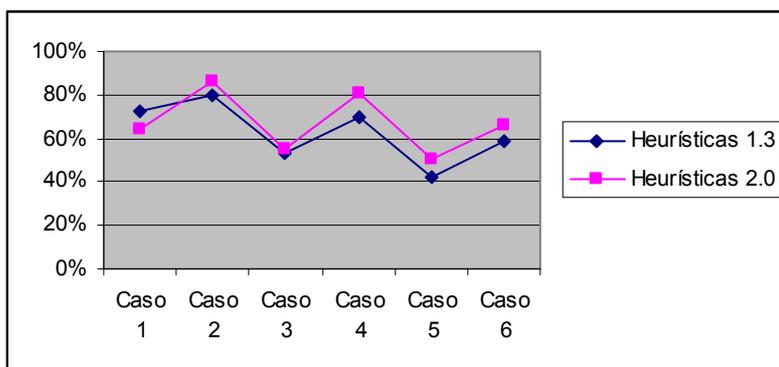


Gráfico 8: Percentual de sumários do Corpus UNU que preservam a textualidade

Já para o Corpus Theses esses resultados foram melhores: os percentuais para foco e textualidade aumentaram em todos os casos na versão 2.0 de heurísticas, ficando acima de 60%. Especialmente, o percentual de sumários que preservam totalmente o foco aumentou consideravelmente, principalmente nos casos 3, 5 e 6.

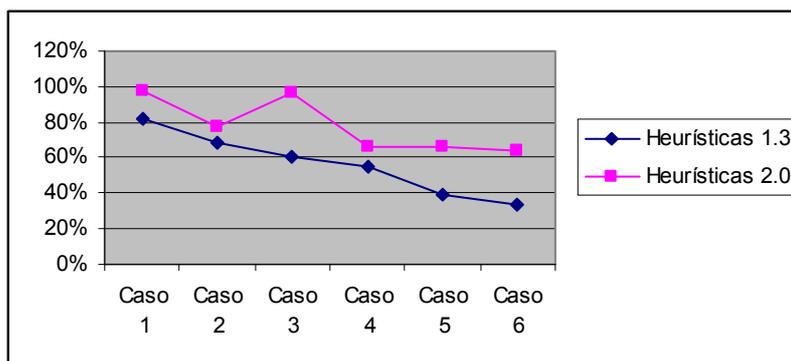


Gráfico 9: Percentual de sumários do Corpus Theses que preservam o foco

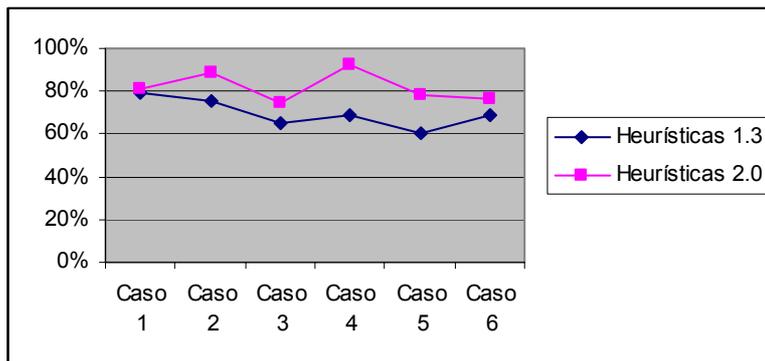


Gráfico 10: Percentual de sumários do Corpus Theses que preservam a textualidade

Um exemplo de aumento de qualidade pode ser observado ao sumarizar a sentença [16] “*In this scope, since diverse areas benefit from the utilization of Internet and some relevant attributes depend on the application domain, the educational area was focalized in this work.*” do Corpus Theses. Utilizando a versão 1.3 do UNLSumm, ela foi sumarizada no Caso 5 por 7 heurísticas - seis do grupo A e uma do Grupo B - gerando o sumário UNL apresentado no Exemplo 32 (RBs tracejadas indicam as RBs excluídas), que poderia ser decodificada como “*Since areas benefit and attributes depend on the domain, the area was focalized*”. Podemos observar que esse sumário, além de não ser bem formado, também não preserva a idéia principal da sentença-fonte, pois não preserva a informação sobre a área focalizada no trabalho. Esse sumário, portanto, foi avaliado com o *score* 1.

Sumarizando essa mesma sentença com a versão 2.0 do UNLSumm, o sumário UNL produzido (Exemplo 33) pela aplicação de 4 heurísticas (três do Grupo A e uma do Grupo B) tem como possível decodificação a sentença “*The educational area was focalized in this work*”. Diferentemente do caso anterior, este sumário é bem formado e mantém a idéia central da sentença-fonte e, portanto, foi avaliado com o *score* 6.

```
[S]
rsn(:04,:01)
and:01(:02,:03)
qua:02(area.@pl,many)
ben:02(benefit.@entry,area.@pl)
src:02(benefit.@entry,Internet.@def)
agt:03(depend,attribute.@pl)
mod:03(attribute.@pl,relevant)
qua:03(attribute.@pl,some)
obj:03(depend,domain.@def)
mod:03(domain.@def,application)
obj:04(focalize.@past,area.@def)
mod:04(area.@def,educational)
plc:04(focalize.@past,work)
mod:04(work,this)
[/S]
```

Exemplo 32: Possível sumário UNL de [16] utilizando o UNLSumm 1.3

```
[S]
rsn(:04,:01)
and:01(:02,:03)
qua:02(area.@pl,many)
ben:02(benefit.@entry,area.@pl)
src:02(benefit.@entry,Internet.@def)
agt:03(depend,attribute.@pl)
mod:03(attribute.@pl,relevant)
qua:03(attribute.@pl,some)
obj:03(depend,domain.@def)
mod:03(domain.@def,application)
obj:04(focalize.@past,area.@def)
mod:04(area.@def,educational)
plc:04(focalize.@past,work)
mod:04(work,this)
[/S]
```

Exemplo 33: Possível sumário UNL de [16] utilizando o UNLSumm 2.0

7 Discussão Final

O principal resultado do estudo de casos foi a exclusão dos subgrupos de heurísticas relacionados aos RLs ‘mod’ e ‘aoj’. No entanto, o estudo também mostrou que outros subgrupos de heurísticas produziram resultados ruins em alguns casos, devendo, portanto, ser melhor analisados. Recorrendo, então, às sentenças correspondentes, vimos que isso ocorre devido à diversidade de uso da interlíngua UNL na codificação. Como a especificação dessa linguagem é flexível, ela pode ser interpretada de formas diferentes por diferentes pessoas. Isso foi verificado especialmente no Corpus UNU, no qual um mesmo tipo de informação foi representado por RLs distintos em sentenças diferentes. A interferência da consistência do corpus de textos UNL na sumarização pode ser comprovada ao verificarmos que o Corpus Theses, codificado manualmente e de forma consistente por nós mesmos, apresentou melhores sumários em todos os casos.

Esses resultados indicam que as heurísticas de sumarização levarão a melhores resultados quando os textos UNL a serem sumarizados forem codificados de forma consistente. Entretanto, este já é um problema de se garantir a qualidade do dado de entrada do UNLSumm, de responsabilidade do processo prévio de codificação da língua natural de origem em UNL. Embora esteja fora do escopo do UNLSumm, para garantir a qualidade dos sumários UNL deveríamos buscar sua robustez. Para isso, uma proposta é utilizar apenas heurísticas que comprovadamente identifiquem e excluam somente informações supérfluas. Outra proposta é, futuramente, considerar modelos dependentes de língua natural e, assim, criar versões de heurísticas que sejam dependentes de língua ou da própria cultura. Nesse sentido, considerando que outros RLs além de ‘mod’ e ‘aoj’ foram apontados como tendo aplicabilidade 0 nos corpora deste estudo, por exemplos, os RLs ‘src’, ‘met’ e ‘nam’ para o Corpus UNU (vide seção 4.1.4), seria interessante, ao utilizar outros dados de teste, verificar sua aplicabilidade, para refinar o conjunto de heurísticas do UNLSumm, no sentido de bem caracterizar as construções conceituais que mais expressivamente reflitam informações supérfluas e, dessa forma, tornar o UNLSumm não só de uso geral mas limitando-o, para usos mais específicos.

Em relação às taxas de compressão obtidas, verificamos que, quando se aplicam heurísticas de ambos os grupos, elas tendem a se situar próximas aos 40% e, de um modo geral, acima de 45% apresentam baixos resultados (i.e., sumários de baixa qualidade, em relação às hipóteses deste trabalho). Assim, esses resultados indicam que, para a sumarização automática intra-sentencial, a taxa de compressão média deve ficar em torno de 40%.

Essas taxas podem ser consideradas limitantes em um sistema de sumarização automática com abordagem profunda de reescrita do texto, que, em geral, reduzem de 70 a 95% o texto original (Mani, 2001). Entretanto, considerando que contemplamos somente a exclusão de segmentos sentenciais, a taxa usual dos sistemas de sumarização automática intra-sentencial é reduzida para, aproximadamente, 42% (veja, p. ex., Jing, 2000). Assim, adotaremos como limite máximo para a taxa de compressão, no UNLSumm, o valor correspondente a 45% do texto-fonte, para o qual este estudo de casos já evidenciou a obtenção de resultados satisfatórios.

É importante ressaltar que, tanto em relação à taxa de compressão quanto em relação à abordagem utilizada, os resultados não podem ser considerados conclusivos, porque: a) foram utilizados somente dois corpora de textos, muito embora eles totalizem, para o UNLSumm, 4336 relações binárias envolvidas em decisões de sumarização; b) foram considerados domínios e gêneros restritos, para os dados e c) o estudo de casos

constituiu, na verdade, de uma primeira abordagem visando garantir a confiabilidade das heurísticas propostas, indicando alterações já incluídas na versão 2.0 do UNLSumm, as quais, entretanto, ainda não foram apropriadamente testadas com outros dados. Apesar disso, o estudo de casos aqui relatado apontou diversas contribuições, sendo a mais imediata a produção de outra versão do repositório de heurísticas.

Em relação à taxa de compressão, pretendemos explorar melhor o patamar de 45% de compressão do texto-fonte, permitindo, p. ex., que o usuário customize o sistema para gerar sumários de acordo com uma taxa dada, a exemplo do que se pode fazer com o *Auto Summarizer* do *Microsoft Word*. Além disso, se esse usuário for especialista no UNLSumm (como, normalmente, será o caso, nesta etapa de desenvolvimento), ele poderá indicar, ainda, qual o grupo de heurísticas que pretende explorar para, assim, analisar de perto a influência de cada estratégia na sumarização intra-sentencial de códigos UNL, a exemplo do que fizemos neste estudo, ao adotar a hipótese 5 (de interdependência entre as heurísticas do Grupo A e do Grupo B).

Vale notar que cada opção dessas implicará em uma clara interação do usuário com o sistema, assim como a produção de novos resultados que deverão, então, ser avaliados. Neste caso, as próprias métricas de avaliação delineadas neste relatório (vide seção 3.1), de textualidade e preservação da idéia central, assim como de utilizar ambas como pontos de decisão conjugados para o intervalo de medidas de avaliação, poderão ser revistas. Uma possível revisão, neste caso, já necessária, seria em relação ao peso atribuído a cada um desses pontos: a tendência de preservar a idéia principal, em detrimento da textualidade, deverá ser eliminada, fazendo com que o equilíbrio entre ambos os pontos de decisão reflita a qualidade de um texto quando, por exemplo, ele não for considerado tão dependente de seu correspondente texto-fonte (caso em que a textualidade assume maior importância).

Referências Bibliográficas

- Feltrim, V.D.; Nunes, M.G.V.; Aluísio, S.M. (2001). *Um corpus de textos científicos em Português para a análise da Estrutura Esquemática*. Tech. Rep. NILC-TR-01-04. São Carlos, SP, Julho.
- Jing, H. (2000). "Sentence Reduction for Automatic Text Summarization". In *Proceedings of the 6th Applied Natural Language Processing Conference and the First Meeting of the North American Chapter of the Association for Computational Linguistics (ANLP-NAACL'2000)*, pp. 310-315. New Brunswick, New Jersey.
- Mani, I. (2001). *Automatic Summarization*. John Benjamin's Publishing Company. USA.
- Martins, C. B. (2001). *A Sumarização de Textos UNL*. Monografia de Qualificação. PPG-CC, DC-UFSCar. São Carlos, SP.
- Martins, C. B. e Rino, L. H. M. (2002). *Heurísticas de Poda de Sentenças para a Sumarização Automática de Textos UNL*. Série de Relatórios do NILC, NILC-TR-02-01. São Carlos, SP, Março, 38p.
- Rino, L.H.M. (1996). *Modelagem de Discurso para o Tratamento da Concisão e Preservação da Idéia Central na Geração de Textos*. Tese de Doutorado. Universidade de São Paulo, Brasil.
- Uchida, H. (1997). *DeConverter Specification, Version 1.0*. Tech. Rep. UNL-TR1997-010, UNU/IAS/UNL Center, Tokyo, Japan.
- Uchida, H. (2000). *Universal Networking Language: An Electronic Language for Communication, Understanding and Collaboration*. UNL Center, IAS/UNU, Tokyo (também disponível no site www.unl.ias.unu.edu).
- Vallduvi, E. (1993). *Information Packaging: A survey*. Research Paper HCRC/RP-44. HCRC Publications. University of Edinburgh. Scotland, UK.
- White, J., Doyon, J. and Talbott, S. (2000). Task tolerance of MT output in Integrated Text Processes. *Proceedings of the Embedded MT Systems Workshop*, pp. 9-16 (NAACL-ANLP 2000 Workshop). Seattle, WA.

Anexo 1

Versão 1.3 das Heurísticas de Sumarização

RL	GRUPO A	GRUPO B
met	<p>HA.1.1. Excluir met(a,b) - se b \notin outras RBs e não existe obj(a,x).</p>	<p>HB.1.1. Excluir met(a,b) obj(b, c) mod(c, d) – se b, c,d \notin RBs fora do subgrupo. HB.1.2. Excluir met(a,b) obj(a,c) aoj(d,b) – se b,c e d \notin RBs fora do subgrupo. HB.1.3. Excluir met(x,b) obj(b,a) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.1.4. Excluir met(x, b) mod(b,c) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.1.5. Excluir met(a,b) aoj(c,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.1.6. Excluir met(a,b) + {RBs com escopo ‘b’} – se eh_escopo(b). HB.1.7. Excluir met(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo.</p>
man	<p>HA.2.1. Excluir man(a,b) - se b \notin outras RBs e existe aoj(a,x). HA.2.2. Excluir man(a,b) – se b \notin outras RBs e existe obj(a,x). HA.2.3. Excluir man(a,b) – se b \notin outras RBs e existe mod(x,a). HA.2.4. Excluir man(a,b) - se b \notin outras RBs e existe ins(a,x). HA.2.5. Excluir man(a,b) – se b \notin outras RBs e existe agt(a,x). HA.2.6. Excluir man(a,b) – se b \notin outras RBs e existe gol(a,x). HA.2.7. Excluir man(a,b) – se b \notin outras RBs e existe mod(a,x). HA.2.8. Excluir man(a,b) – se b \notin outras RBs e existe bem(a,x).</p>	<p>HB.2.1. Excluir man(a,b) obj(b,c) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.2.2. Excluir man(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.2.3. Excluir man(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>
lpl	<p>HA.3.1. Excluir lpl(a,b) – se b \notin outras RBs.</p>	<p>HB.3.1. Excluir lpl(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.3.2. Excluir lpl(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>
ppl	<p>HA.3.2. Excluir ppl(a,b) – se b \notin outras RBs.</p>	<p>HB.3.3. Excluir ppl(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.3.4. Excluir ppl(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>
plc	<p>HA.3.3. Excluir plc(a,b) – se b \notin outras RBs.</p>	<p>HB.3.5. Excluir plc(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.3.6. Excluir plc(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>
aoj	<p>HA.4.1. Excluir aoj(a,b) – se a \notin outras RBs e existe obj(x,b). HA.4.2. Excluir ben(a,b) – se a \notin outras RBs e existe ben(x,b). HA.4.3. Excluir mod(a,b) – se a \notin outras RBs e existe mod(b,x).</p>	<p>HB.4.1. Excluir aoj(a,b) obj(a,c) – se b, c \notin RBs fora do subgrupo. HB.4.2. Excluir aoj(a,b) obj(a,c) – se existe agt(d,b) e obj(d,e). HB.4.3. Excluir aoj(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo.</p>
tim	<p>HA.5.1. Excluir tim(a,b) – se b \notin outras RBs e</p>	<p>HB.5.1. Excluir tim(a,b) + {RBs \in subgrupo</p>

	<p>existe $\text{man}(a,x)$.</p> <p>HA.5.2. Excluir $\text{tim}(a,b)$ – se $b \notin$ outras RBs e existe $\text{agt}(a,x)$.</p> <p>HA.5.3. Excluir $\text{tim}(a,b)$ – se $b \notin$ outras RBs e existe $\text{obj}(a,x)$.</p> <p>HA.5.4. Excluir $\text{tim}(a,b)$ – se $b \notin$ outras RBs e existe $\text{mod}(a,x)$.</p> <p>HA.5.5. Excluir $\text{tim}(a,b)$ – se $b \notin$ outras RBs e existe $\text{aoj}(a,x)$.</p> <p>HA.5.6. Excluir $\text{tim}(a,b)$ – se $b \notin$ outras RBs.</p>	<p>$S1\}$ – se $UWs \in S1 \notin$ RBs fora do subgrupo.</p> <p>HB.5.2. Excluir $\text{tim}(a,b)$ + RBs com escopo 'b' – se $\text{eh_escopo}(b)$.</p>
pos	HA.6.1. Excluir $\text{pos}(a,b)$ – se $b \notin$ outras RBs.	-
mod	<p>HA.7.1. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{and}(a,x)$ ou $\text{and}(x,a)$.</p> <p>HA.7.2. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{aoj}(a,x)$.</p> <p>HA.7.3. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{ben}(x,a)$.</p> <p>HA.7.4. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{lpl}(x, a)$.</p> <p>HA.7.5. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{mod}(x,a)$.</p> <p>HA.7.6. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{mod}(a,x)$.</p> <p>HA.7.7. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{obj}(x,a)$.</p> <p>HA.7.8. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{obj}(a,x)$.</p> <p>HA.7.9. Excluir $\text{mod}(a, b)$ – se $b \notin$ outras RBs e existe $\text{pof}(a, x)$.</p> <p>HA.7.10. Excluir $\text{mod}(a,b)$ - se $b \notin$ outras RBs e existe $\text{qua}(a,x)$.</p> <p>HA.7.11. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{aoj}(x,a)$.</p> <p>HA.7.12. Excluir $\text{mod}(a,b)$ – se \notin outras RBs e existe $\text{agt}(x,a)$.</p> <p>HA.7.13. Excluir $\text{mod}(a,b)$ – se $b \notin$ outras RBs e existe $\text{or}(x,a)$ ou $\text{or}(a,x)$.</p>	<p>HB.7.1. Excluir $\text{mod}(a,b)$ $\text{aoj}(c,b)$ – se $b,c \notin$ RBs fora do subgrupo.</p> <p>HB.7.2. Excluir $\text{mod}(a,b)$ $\text{and}(b,c)$ - se $b,c \notin$ RBs fora do subgrupo.</p> <p>HB.7.3. Excluir $\text{mod}(a,b)$ + RBs com escopo 'b' – se $\text{eh_escopo}(b)$.</p> <p>HB.7.4. Excluir $\text{mod}(a,b)$ + $\{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin$ RBs fora do subgrupo.</p>
nam	HA.8.1. Excluir $\text{nam}(a,b)$ – se $b \notin$ outras RBs.	-
src	HA.9.1. Excluir $\text{src}(a,b)$ – se $b \notin$ outras RBs e existe $\text{obj}(x,a)$.	<p>HB.9.1. Excluir $\text{src}(a,b)$ + $\{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin$ RBs fora do subgrupo.</p> <p>HB.9.2. Excluir $\text{src}(a,b)$ $\text{tim}:b(c,d)$ – se $c,d \notin$ RBs fora do subgrupo.</p> <p>HB.9.3. Excluir $\text{src}(a,b)$ + RBs com escopo b – se $\text{eh_escopo}(b)$.</p>
qua	HA.10.1. Excluir $\text{qua}(a,b)$ – se $b \notin$ outras RBs.	-
cnt	HA.11.1. Excluir $\text{cnt}(a,b)$ – se $b \notin$ outras RBs.	-
scn		<p>HB.12.1. Excluir $\text{scn}(a,b)$ + $\{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin$ RBs fora do subgrupo.</p> <p>HB.12.2. Excluir $\text{scn}(a,b)$ + RBs com escopo b – se $\text{eh_escopo}(b)$.</p>
rsn	-	<p>HB.13.1. Excluir $\text{rsn}(a,b)$ + $\{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin$ RBs fora do subgrupo.</p> <p>HB.13.3. Excluir $\text{rsn}(a,b)$ + RBs com escopo b – se $\text{eh_escopo}(b)$.</p>

cau	-	<p>HB.13.2. Excluir $\text{cau}(a,b) + \{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin \text{RBs}$ fora do subgrupo.</p> <p>HB.13.4. Excluir $\text{cau}(a,b) + \text{RBs}$ com escopo b – se $\text{eh_escopo}(b)$.</p>
pur	-	<p>HB.14.1. Excluir $\text{pur}(a,b) \text{ obj}(b,c) + \{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin \text{RBs}$ fora do subgrupo.</p> <p>HB.14.2. Excluir $\text{pur}(a,b) + \{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin \text{RBs}$ fora do subgrupo.</p> <p>HB.14.3. Excluir $\text{pur}(a,b) + \text{RBs}$ com escopo b – se $\text{eh_escopo}(b)$.</p>
dur	-	<p>HB.15.1. Excluir $\text{dur}(a,b) \text{ mod}(b,c) + \{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin \text{RBs}$ fora do subgrupo.</p> <p>HB.15.2. Excluir $\text{dur}(a,b) + \{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin \text{RBs}$ fora do subgrupo.</p> <p>HB.15.3. Excluir $\text{dur}(a,b) + \text{RBs}$ com escopo b – se $\text{eh_escopo}(b)$.</p>
con	-	<p>HB.16.1. Excluir $\text{con}(a,b) + \{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin \text{RBs}$ fora do subgrupo.</p> <p>HB.16.2. Excluir $\text{con}(a,b) + \text{RBs}$ com escopo b – se $\text{eh_escopo}(b)$.</p>
agt	-	<p>HB.17.1. Excluir $\text{agt}(a,b) \text{ obj}(a,c) + \{\text{RBs} \in \text{subgrupo } S1\}$ – se existe $\text{agt}(x,b)$ e $UWs \in S1 \notin \text{RBs}$ fora do subgrupo.</p>
ben	-	<p>HB.18.1. Excluir $\text{ben}(a,b) + \{\text{RBs} \in \text{subgrupo } S1\}$ – se $UWs \in S1 \notin \text{RBs}$ fora do subgrupo.</p> <p>HB.18.2. Excluir $\text{ben}(a,b) + \text{RBs}$ com escopo b – se $\text{eh_escopo}(b)$.</p>

Anexo 2

Versão 2.0 das Heurísticas de Sumarização

RL	GRUPO A	GRUPO B
met	<p>HA.1.1 - Excluir met(a,b) - se b \notin outras RBs e não existe obj(a,x).</p>	<p>HB.1.1. Excluir met(a,b) obj(b, c) mod(c, d) – se b, c,d \notin RBs fora do subgrupo. HB.1.2. Excluir met(a,b) obj(a,c) aoj(d,b) – se b,c e d \notin RBs fora do subgrupo. HB.1.3. Excluir met(x,b) obj(b,a) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.1.4. Excluir met(x, b) mod(b,c) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.1.5. Excluir met(a,b) aoj(c,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.1.6. Excluir met(a,b) + {RBs com escopo ‘b’} – se eh_escopo(b). HB.1.7. Excluir met(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo.</p>
man	<p>HA.2.1. Excluir man(a,b) - se b \notin outras RBs e existe aoj(a,x). HA.2.2. Excluir man(a,b) – se b \notin outras RBs e existe obj(a,x). HA.2.3. Excluir man(a,b) – se b \notin outras RBs e existe mod(x,a). HA.2.4. Excluir man(a,b) - se b \notin outras RBs e existe ins(a,x). HA.2.5. Excluir man(a,b) – se b \notin outras RBs e existe agt(a,x). HA.2.6. Excluir man(a,b) – se b \notin outras RBs e existe gol(a,x). HA.2.7. Excluir man(a,b) – se b \notin outras RBs e existe mod(a,x). HA.2.8. Excluir man(a,b) – se b \notin outras RBs e existe ben(a,x).</p>	<p>HB.2.1. Excluir man(a,b) obj(b,c) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.2.2. Excluir man(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.2.3. Excluir man(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>
lpl	<p>HA.3.1. Excluir lpl(a,b) – se b \notin outras RBs.</p>	<p>HB.3.1. Excluir lpl(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.3.2. Excluir lpl(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>
ppl	<p>HA.3.2. Excluir ppl(a,b) – se b \notin outras RBs.</p>	<p>HB.3.3. Excluir ppl(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.3.4. Excluir ppl(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>
plc	<p>HA.3.3. Excluir plc(a,b) – se b \notin outras RBs.</p>	<p>HB.3.5. Excluir plc(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.3.6. Excluir plc(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>
tim	<p>HA.4.1. Excluir tim(a,b) – se b \notin outras RBs e existe man(a,x) HA.4.2. Excluir tim(a,b) – se b \notin outras RBs e existe agt(a,x) HA.4.3. Excluir tim(a,b) – se b \notin outras RBs e existe obj(a,x) HA.4.4. Excluir tim(a,b) – se b \notin outras RBs</p>	<p>HB.4.1. Excluir tim(a,b) + {RBs \in subgrupo S1} – se UWs \in S1 \notin RBs fora do subgrupo. HB.4.2. Excluir tim(a,b) + RBs com escopo ‘b’ – se eh_escopo(b).</p>

	e existe mod(a,x) HA.4.5. Excluir tim(a,b) – se b ∉ outras RBs e existe aoj(a,x) HA.4.6. Excluir tim(a,b) – se b ∉ outras RBs	
pos	HA.5.1. Excluir pos(a,b) – se b ∉ outras RBs.	-
nam	HA.6.1. Excluir nam(a,b) – se b ∉ outras RBs	-
src	HA.7.1. Excluir src(a,b) – se b ∉ outras RBs e existe obj(x,a).	HB.7.1. Excluir src(a,b) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.7.2. Excluir src(a,b) tim:b(c,d) – se c,d ∉ RBs fora do subgrupo. HB.7.3. Excluir src(a,b) + RBs com escopo b – se eh_escopo(b)
qua	HA.8.1. Excluir qua(a,b) – se b ∉ outras RBs.	-
cnt	HA.9.1. Excluir cnt(a,b) – se b ∉ outras RBs.	
scn	HA.10.1. Excluir scn(a,b) – se b ∉ outras RBs e existe obj(x,a).	HB.10.1. Excluir scn(a,b) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.10.2. Excluir scn(a,b) + RBs com escopo b – se eh_escopo(b).
rsn	-	HB.11.1. Excluir rsn(a,b) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.11.3. Excluir rsn(a,b) + RBs com escopo b – se eh_escopo(b)
cau	-	HB.11.2. Excluir cau(a,b) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.11.4. Excluir cau(a,b) + RBs com escopo b – se eh_escopo(b).
pur	-	HB.12.1. Excluir pur(a,b) obj(b,c) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.12.2. Excluir pur(a,b) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.12.3. Excluir pur(a,b) + RBs com escopo b – se eh_escopo(b).
dur	-	HB.13.1. Excluir dur(a,b) mod(b,c) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.13.2. Excluir dur(a,b) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.13.3. Excluir dur(a,b) + RBs com escopo b – se eh_escopo(b).
con	-	HB.14.1. Excluir con(a,b) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.14.2. Excluir con(a,b) + RBs com escopo b – se eh_escopo(b).
agt	-	HB.15.1. Excluir agt(a,b) obj(a,c) – se existe agt(x,b) e c ∉ RBs fora do subgrupo.
ben	-	HB.16.1. Excluir ben(a,b) + {RBs ∈ subgrupo S1} – se UWs ∈ S1 ∉ RBs fora do subgrupo. HB.16.2. Excluir ben(a,b) + RBs com escopo b – se eh_escopo(b).