

Sumarização Automática: Métodos e Avaliação

Thiago A. S. Pardo

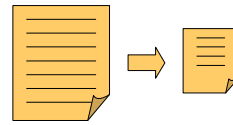


I Escola de Verão
Linguateca
2006



Sumarização

- Produção de uma versão mais curta de um texto-fonte: seu sumário



- Sumário, resumo
 - Extrato e *abstract*

Sumarização

- Permeia o dia a dia das pessoas
 - Sinopse de novelas
 - Resumo de notícias
 - Resenhas de livros e filmes
 - Abstracts de artigos científicos
 - Passagens de páginas da internet

[Forum-[lp](#)] [Linguateca - Escola de Verão](#)

A [Linguateca](#) tem o prazer de anunciar a Primeira [Escola de Verão](#), que terá lugar na Universidade do Porto, Portugal, de 10 a 14 de Julho de 2006. ...
<https://mail.di.fct.unl.pt/pipermail/forum-lp/2006-February/000092.html> - 4k -
[Em cache](#) - [Páginas Semelhantes](#)

Sumarização

- Motivações (humanas)
 - Acesso rápido à informação (*aboutness*)
 - Auxílio à tomada de decisões
 - Comprar um livro, alugar um filme, ler uma tese?
 - Acessar uma página da internet
 - Incapacidade de se absorver toda a informação disponível
 - Estudo de Berkeley (2003)
 - 5 milhões de terabytes de nova informação (filme, meio magnético, impressa, *on-line*, etc.)
 - *Web*: 170 terabytes
 - Dobro do produzido em 1999
 - Aumento de 30% por ano



Sumarização

- Normalmente, resumos são textos
 - Devem apresentar as mesmas características que atribuem 'textualidade' (tessitura) a um texto.
 - Coerência e coesão
 - Boa progressão temática
 - Gramaticalidade
 - Legibilidade
 - Etc.
 - Além de
 - Informação relevante



Sumarização: conceitos

- Taxa de compressão
 - O quanto "enxugar" o texto
 - Dependente da aplicação do sumário
- Informação
 - Essencial, complementar, supérflua
 - Dependente da audiência
- Idéia principal
 - Comunicada pelo escritor
 - Entendida pelo leitor



Sumarização: conceitos

- Tipos de sumário
 - Informativo (autocontido)
 - Indicativo (indexador)
 - Crítico
- Modo de produção
 - Extratos
 - *Abstracts*

Sumarização

- Fatores que influenciam
 - Audiência: genérica ou especializada
 - Objetivo do sumário: substituir o texto-fonte, indexar, criticar
 - Fluência: textual ou fragmentado
 - Fonte: mono ou multidocumento



Sumarização

- Humana
 - Grande variedade de sumários para um mesmo texto
 - Processo quase intuitivo
- Computacional?
 - Como simular a habilidade humana?

Sumarização

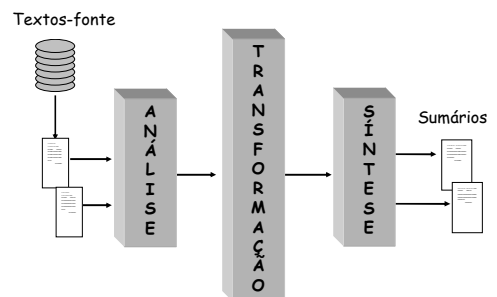
- Motivações (lingüístico-computacionais)
 - Acesso somente à informação relevante
 - Recuperação de informação
 - Extração de informação
 - Categorização textual
 - Perguntas e respostas
 - Produção de sumários úteis aos humanos
 - Desafio: o computador deve 'entender' a língua
 - Envolve todas as questões mais complexas de Processamento de Língua Natural (PLN)
 - Interpretação textual
 - Geração textual
 - Avaliação



Sumarização Automática

- Financiamento massivo
 - EUA/DARPA, Comunidade Européia, Pacific Rim
 - Interesses governamentais e comerciais
- História
 - Primeiro sistema na década de 50
 - Acompanhou a história da IA: 'morte' e 'renascimento' da pesquisa
 - De extratos para *abstracts*
 - Hoje
 - Conferências dedicadas ao tema
 - Da sumarização para as aplicações

Sumarização Automática



Sumarização Automática

- Análise: interpretação do texto-fonte e representação do conteúdo do texto-fonte
- Transformação: produção da representação do conteúdo do sumário
- Síntese: geração do sumário a partir de sua representação de conteúdo

Operações de sumarização



- Seleção/eliminação
 - Seleção do que é relevante ou exclusão do que é irrelevante
- Agregação
 - Associação (*merge*) de informações diversas
- Generalização/substituição
 - Substituição de informações específicas por informação mais geral

Sumarização Automática

- Duas abordagens principais
 - Superficial: estatística, empírica
 - Processo menos complexo
 - Robustez
 - Resultados piores
 - Profunda: lingüística, fundamental
 - Processo mais complexo
 - Especificidade para alguns domínios
 - Resultados melhores
- Abordagens híbridas

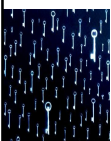


Sumarização Automática

- Abordagens superficiais
 - Extratos
- Abordagens profundas
 - *Abstracts*
 - Operações de sumarização

Abordagem superficial

- Método das palavras-chave
 - Luhn, 1958; Edmundson, 1969; Black e Johnson, 1988
 - O escritor do texto utiliza palavras-chave para expressar a idéia principal
 - As palavras-chave se repetem no decorrer do texto
 - Seleção de sentenças que contêm palavras-chave para compor o extrato



Abordagem superficial

- Método da localização
 - Baxendale (1958)
 - Sentenças importantes ocorrem em lugares mais proeminentes do texto
 - Início e fim de parágrafo/texto



Abordagem superficial

o Método das palavras e frases indicativas

- Paice (1981)
- Seleção de sentenças cujo conteúdo é sinalizado como relevante por palavras e frases indicativas
 - o Artigo científico: "o objetivo deste trabalho..."
 - o Esporte: "resultado", "placar"



Abordagem superficial

o Método relacional

- Skorochodko (1971)
- As sentenças mais importantes são aquelas altamente relacionadas às outras (co-ocorrência de palavras/conceitos)



Abordagem superficial

o Mineração de textos

- Larocca Neto et al. (2000)
- TF-ISF (*Term Frequency – Inverse Sentence Frequency*)
 - o Quanto mais representativas as palavras de uma sentença, mais importante ela é no texto



Abordagem superficial

o Método da idéia principal

- Pardo et al. (2003)
- Há uma sentença identificável no texto que expressa sua idéia principal
 - o O sumário é construído a partir desta sentença



Abordagem profunda

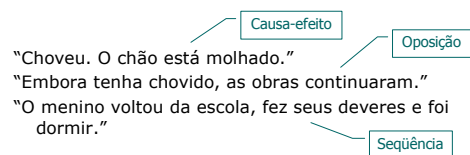
o Conhecimento lingüístico e extralingüístico

- Regras de interpretação e geração textual
- Modelos de língua
- *Wordnets*
- Gramáticas
- Discurso
 - o Identificação do que é relevante no contexto
 - o Diversas teorias discursivas

Discurso

o Um texto é mais do que uma simples seqüência de sentenças justapostas

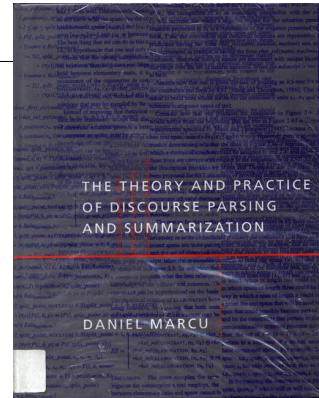
- Estrutura altamente elaborada
 - o Coerência/sentido



Análise discursiva

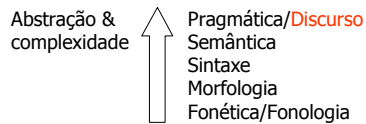
- Análise discursiva relacional (Moore e Pollack, 1992; Moser e Moore, 1996)
 - Lingüística textual
- Análise discursiva automática: identificação automática da estrutura subjacente ao texto

Discurso vs. discurso



Discurso

- Níveis de conhecimento em PLN



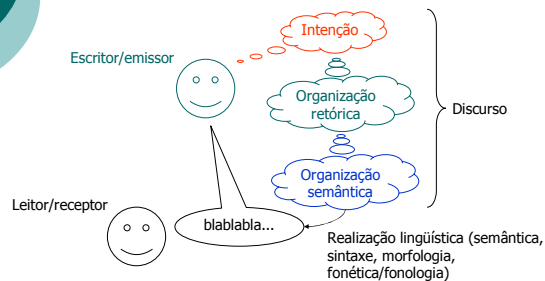
- Envolve a situação de comunicação (Koch e Travaglia, 2002): escritor e leitor

Teorias discursivas

- Grosz e Sidner (1986): intenções
- Mann e Thompson (1987): retórica
- Jordan (1992) e Kehler (2002): semântica
- Moore e Pollack (1992), Moore e Paris (1993), Korelsky e Kittredge (1993), Moser e Moore (1996), Rino (1996) e Marcu (1999, 2000), entre outros: mapeamentos entre os níveis do discurso

Níveis do discurso

- Intenções, retórica e semântica (Moore e Paris, 1993; Korelsky e Kittredge, 1993)



Rhetorical Structure Theory – RST

Mann e Thompson, 1987

- Retórica: parte “palpável” da pragmática (Hovy, 1988)
- Meio pelo qual um texto é organizado para satisfazer um objetivo comunicativo
 - Intenção
- Organização funcional do texto
 - Função de suas partes para o sucesso da comunicação

Rhetorical Structure Theory – RST

Mann e Thompson, 1987

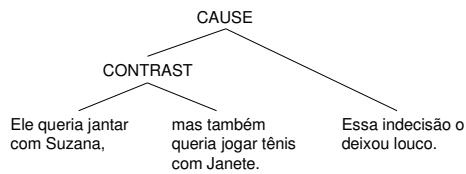
o Estrutura hierárquica do texto

- Relações retóricas entre proposições (unidades de conteúdo) expressas no texto
 - o Em geral, proposições simples são expressas por orações
- Núcleos e satélites
 - o Relações mononucleares e multinucleares
- Relações intencionais e informativas
 - o Intencionais: alteram a inclinação do leitor para algo
 - o Informativas: informam o leitor sobre algo

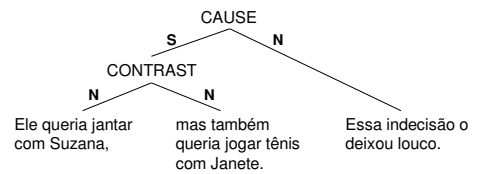
RST: exemplo

Ele queria jantar com Suzana, mas também queria jogar tênis com Janete. Essa indecisão o deixou louco.

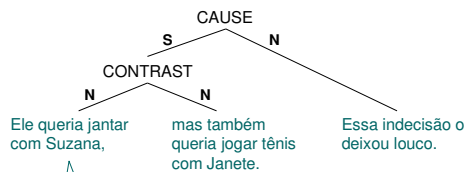
RST: exemplo



RST: exemplo

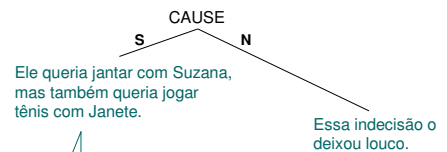


RST: exemplo



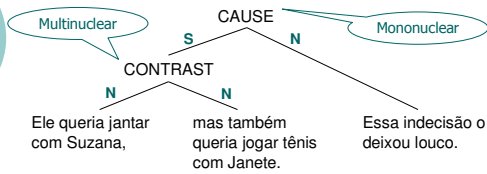
Segmentos que expressam proposições

RST: exemplo

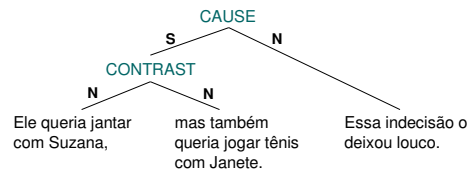


Segmentos que expressam proposições

RST: exemplo

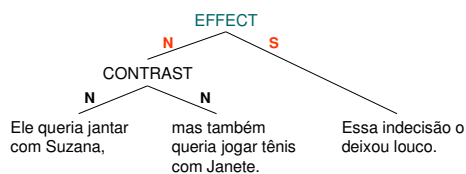


RST: exemplo



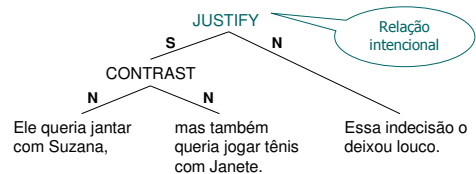
Relações informativas

RST: exemplo



Relações informativas

RST: exemplo



- Como decidir por uma relação?

Determinação de uma relação entre duas proposições

- Definição com 4 campos
 - Restrições sobre o núcleo (N)
 - Restrições sobre o satélite (S)
 - Restrições sobre o núcleo e o satélite
 - Efeito
- Verificação das restrições para as proposições envolvidas
 - Escrita ou leitura de um texto

Definição de relações

Relação: JUSTIFY

Restrições sobre N: não há

Restrições sobre S: não há

Restrições sobre N+S: a compreensão de S pelo leitor aumenta sua prontidão para aceitar o direito do escritor de apresentar N

Efeito: a prontidão do leitor para aceitar o direito do escritor de apresentar N aumenta

Definição de relações

Relação: CAUSE

- Restrições sobre N:** apresenta uma ação volitiva ou uma situação que poderia surgir de uma ação volitiva
- Restrições sobre S:** não há
- Restrições sobre N+S:** S apresenta uma situação que pode ter acarretado o fato do agente da ação volitiva em N ter realizado a ação; sem S, o leitor poderia não reconhecer a motivação da ação; N é mais central para a satisfação do objetivo do escritor do que S
- Efeito:** o leitor reconhece que a situação apresentada em S como a causa da ação apresentada em N

Definição de relações

Relação: EVIDENCE

- Restrições sobre N:** o leitor pode não acreditar em N de forma satisfatória para o escritor
- Restrições sobre S:** o leitor acredita em S ou o acha válido
- Restrições sobre N+S:** a compreensão de S pelo leitor aumenta sua crença em N
- Efeito:** a crença do leitor em N aumenta

Definição de relações

Relação: CONCESSION

- Restrições sobre N:** o escritor julga N válido
- Restrições sobre S:** o escritor não afirma que S pode não ser válido
- Restrições sobre N+S:** o escritor mostra uma incompatibilidade aparente ou em potencial entre N e S; o reconhecimento da compatibilidade entre N e S melhora a aceitação de N pelo leitor
- Efeito:** o leitor aceita melhor N

RST

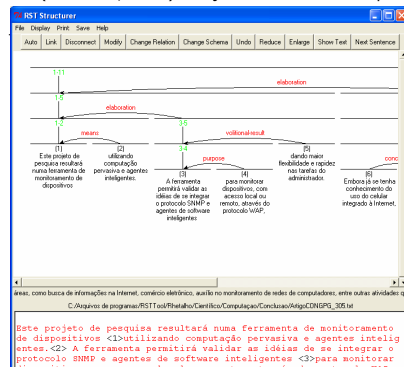
- Subjetividade da análise e ambigüidade
 - Determinação da nuclearidade
 - NS, SN ou NN?
 - Escolha das relações
 - CAUSE, EFFECT ou JUSTIFY?
 - CONCESSION, CONTRAST, ANTI-THESIS ou OTHERWISE?
 - Determinação dos segmentos
 - Orações, sentenças ou parágrafos?
 - Fragmentos?
 - Forma de representação da análise
 - Árvore ou grafo (Wolf e Gibson, 2005)?

RST

- Subjetividade da análise e ambigüidade
 - Originalmente, descrição de textos
 - Posteriormente, manipulação computacional de textos

Ferramentas de suporte

RSTTool (O'Donnel, 1997): edição e facilidades computacionais



The screenshot shows the RSTTool application window. The title bar reads 'RST Structur'. The menu bar includes 'File', 'Display', 'View', 'Save', and 'Help'. The toolbar contains icons for 'Auto', 'Link', 'Disconnect', 'Modify', 'Change Fileset', 'Change Schema', 'Undo', 'Redo', 'Edge', 'Show Text', and 'Next Sentence'. The main text area displays a paragraph with RST annotations: 'elaboration' above the first sentence, 'main' above the second, 'elaboration' above the third, and 'volitional' above the fourth. The text is: 'Este sistema de pesquisa realiza... (1) elaboração... (2) computação... (3) para monitorar... (4) dispositivos... (5) Embody o sistema...'. The status bar at the bottom shows the file path: 'C:\Arquivos de programas\RSTTool\Ferramentas\Cercitico\Computacao\Conclusao\Relato\CONTEPO_305.txt'. Below the window, there is a red text box with the following text: 'Este projeto de pesquisa resultará numa ferramenta de monitoramento de dispositivos <1>utilizando computação pervasiva e agentes inteligentes <2> A ferramenta permitirá validar as ideias de se integrar o protocolo SNMP e agentes de software inteligentes <3>para monitorar...'

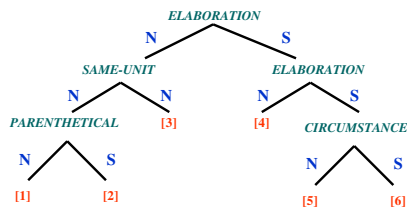
Analisadores discursivos automáticos

- Inglês
 - Marcu (1997, 2000)
 - Corston-Oliver (1998)
 - Schilder (2002)
 - Marcu e Echihabi (2002)
 - Soricut e Marcu (2003)
 - Reitter (2003)
 - Hanneforth et al. (2003)
 - Mahmud e Ramsay (2005)
- Japonês
 - Sumita et al. (1992)
- Português
 - DiZer (Pardo, 2005)

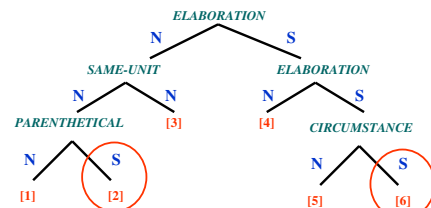
Abordagem profunda

- Idéia básica para sumarização
 - Satélites são informação complementar e, portanto, podem ser eliminados
 - Vários métodos para se escolher que segmentos eliminar
 - Mann e Thompson (1992), Rino (1996), O'Donnel (1997), Marcu (2000)

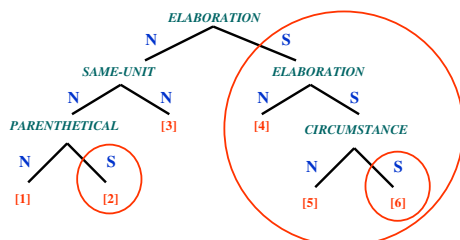
[1] A empresa Produtos Pirata Indústria e Comércio Ltda., de Contagem [2] (na região metropolitana de Belo Horizonte), [3] deverá registrar este ano um crescimento de produtividade nas suas áreas comercial e industrial de 11% e 17%, respectivamente. [4] Os ganhos são atribuídos pela diretoria da fábrica à nova filosofia [5] que vem sendo implantada na empresa desde outubro do ano passado, [6] quando a Pirata se iniciou no Programa Sebrae de Qualidade Total.



[1] A empresa Produtos Pirata Indústria e Comércio Ltda., de Contagem [2] (na região metropolitana de Belo Horizonte), [3] deverá registrar este ano um crescimento de produtividade nas suas áreas comercial e industrial de 11% e 17%, respectivamente. [4] Os ganhos são atribuídos pela diretoria da fábrica à nova filosofia [5] que vem sendo implantada na empresa desde outubro do ano passado, [6] quando a Pirata se iniciou no Programa Sebrae de Qualidade Total.



[1] A empresa Produtos Pirata Indústria e Comércio Ltda., de Contagem [2] (na região metropolitana de Belo Horizonte), [3] deverá registrar este ano um crescimento de produtividade nas suas áreas comercial e industrial de 11% e 17%, respectivamente. [4] Os ganhos são atribuídos pela diretoria da fábrica à nova filosofia [5] que vem sendo implantada na empresa desde outubro do ano passado, [6] quando a Pirata se iniciou no Programa Sebrae de Qualidade Total.



Abordagem profunda

- Mann e Thompson (1992)
 - Eliminação de satélites que não são necessários para que as relações retóricas em foco atinjam seus efeitos pretendidos

Abordagem profunda

- O'Donnel (1997)
 - Cada segmento (núcleo e satélite) tem sua importância determinada em função da profundidade na árvore retórica e da relação a qual pertence

Abordagem profunda

- Marcu (2000)
 - A saliência (profundidade na árvore) de um segmento determina sua importância
 - Quanto mais nuclear, mais importante

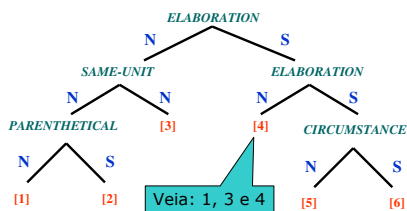
Abordagem profunda

- Rino (1996)
 - A nuclearidade não é suficiente. É necessário considerar o objetivo comunicativo original do texto.

Abordagem profunda

- Teoria das veias (Cristea et al., 1998)
 - Cada segmento da estrutura discursiva contém uma lista de segmentos que possivelmente contém os antecedentes das anáforas que ocorrem
 - Ao se selecionar um segmento para inclusão no sumário, garante-se a inclusão dos segmentos anteriores que possam conter os antecedentes anafóricos

[1] A empresa Produtos Pirata Indústria e Comércio Ltda., de Contagem [2] (na região metropolitana de Belo Horizonte), [3] deverá registrar este ano um crescimento de produtividade nas suas áreas comercial e industrial de 11% e 17%, respectivamente. [4] Os ganhos são atribuídos pela diretoria da fábrica à nova filosofia [5] que vem sendo implantada na empresa desde outubro do ano passado, [6] quando a Pirata se iniciou no Programa Sebrae de Qualidade Total.



Abordagem superficial/profunda

- Aprendizado de máquina
 - Kupiec et al. (1995), Teufel e Moens (1997)
 - Combinação de características sentenciais para julgamento de relevância para compor o sumário
 - Tamanho, posição, número de substantivos, nuclearidade, etc.



Avaliação: como decidir o que é melhor?

- Cenário
 - Diversos métodos
 - Superficiais, profundos e híbridos
 - Variedade de fontes de conhecimento
 - Diversos tipos de sumários
 - Extratos e *abstracts*
 - Genérico ou especializado
 - Textual ou fragmentado
 - Informativo, indicativo ou crítico
 - Mono e multidocumento
 - Taxa de compressão
 - Muitos sumários bons para um mesmo texto



Quesitos avaliáveis



- Desempenho computacional
 - Complexidade do algoritmo, uso de memória, etc.
- Usabilidade
 - Interface, consistência, flexibilidade, etc.
- **Resultados**
 - Qualidade

Forma de avaliação

- **Intrínseca**
 - Qualidade do resultado
 - Quão bom é o sumário?
- **Extrínseca**
 - Aplicação em um contexto
 - O quanto o uso de sumários melhorou a recuperação de informação?

Julgamento humano

- *On-line*
 - Humanos treinados
 - Tempo, dinheiro
 - Questões derivadas da subjetividade
 - Boa descrição da tarefa, concordância
- *Off-line*
 - Reproduzível, rápida e barata

O que se avalia

- *Glass-box*
 - Módulos do sistema
 - Crítica mais elaborada
- *Black-box*
 - Resultado final do sistema
 - O que realmente importa!

Comparação de resultados

- **Avaliação comparativa**
 - Grandes eventos internacionais
 - SUMMAC, DUC
 - *Roadmaps*
- Avaliação autônoma

Como moldar a avaliação?

- Sparck Jones e Galliers (1996)
 - Tão importante quanto a forma de avaliação é saber o que se quer avaliar



Medidas de avaliação intrínseca

- Dois principais aspectos (Mani, 2001)
 - Qualidade textual
 - Informatividade do sumário

Medidas de avaliação intrínseca

- Qualidade
 - Julgamento humano, normalmente
 - Fluência, facilidade de leitura, clareza, legibilidade, concisão
 - Referências anafóricas, explicação para siglas e abreviaturas
 - Integridade das estruturas presentes no texto (listas e tabelas)
 - Coerência e coesão
 - Ortografia e gramática

Medidas de avaliação intrínseca

- Informatividade
 - Cobertura e precisão
 - Informação em comum entre o sumário automático e um sumário ideal
 - Sumário ideal (*gold standard*): humano

$$C = \frac{\text{sentenças Ideal} \cap \text{sentenças Automático}}{\text{sentenças Ideal}}$$

$$P = \frac{\text{sentenças Ideal} \cap \text{sentenças Automático}}{\text{sentenças Automático}}$$



Medidas de avaliação intrínseca

- Medida F
 - Combinação uniforme de cobertura e precisão
 - Medida única de eficiência do sistema: o quão próximo do ideal está

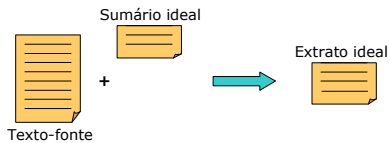
$$F = \frac{2 \times C \times P}{C + P}$$

Medidas de avaliação intrínseca

- Cobertura e precisão
 - Para *abstracts*
 - Intervenção humana
 - Outras medidas de avaliação
 - Para extratos
 - Se sumário ideal for extrato: cômputo automático
 - Caso contrário
 - Produção de um extrato ideal a partir do sumário ideal: manual ou automaticamente
 - Outras medidas de avaliação

Medidas de avaliação intrínseca

- Produção de um extrato ideal a partir de um sumário ideal
 - Selecionam-se as sentenças do texto-fonte que mais se assemelham às sentenças do sumário humano
 - Automaticamente: co-ocorrência e ordenação de palavras



Medidas de avaliação intrínseca

- Medida de utilidade (Radev et al., 2000): extratos
 - Cada sentença do texto-fonte é pontuada (por humanos) de acordo com sua importância
 - O sumário recebe uma nota que corresponde à soma das notas de suas sentenças
- Dificuldades dessa medida
 - Trabalho (des)humano
 - Demanda muito tempo
 - Baixa concordância entre juízes

Medidas de avaliação intrínseca

- Problema das medidas de comparação de sentenças
 - Sentenças diferentes podem transmitir a mesma informação



Medidas de avaliação intrínseca

- Sobreposição de conteúdo
 - Similar à precisão e cobertura, mas considera proposições simples (unidades factuais de informação)
 - Extratos ou *abstracts*: julgamento humano
 - SEE (*Summary Evaluation Environment*)
 - <http://www.isi.edu/~cyl/SEE/>
 - DUC (até 2003)
- *Retention rate* (Mani, 2001)

Medidas de avaliação intrínseca

- Preservação da idéia principal
 - Julgamento humano
 - Cômputo automático: indicação prévia por humanos das partes principais dos textos-fonte



Medidas de avaliação intrínseca

- ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) (Lin e Hovy, 2003)
 - <http://www.isi.edu/~cyl/ROUGE/>
 - Automática, com julgamento próximo do humano
 - Co-ocorrência de n -gramas entre sumário automático e sumário(s) de referência
 - Seqüências de palavras: 1 a 4
- Ceticismo: "medidas automáticas podem ser enganadas"
- Explosão das pesquisas em avaliação em sumarização

Medidas de avaliação intrínseca

- Recentemente...
 - Grafos: informatividade
 - Erkan and Radev (2004); Mihalcea (2005); Santos Jr. et al. (2004)
 - Textos e sumários são representados como grafos e comparados
 - Redes complexas: qualidade
 - Antiqueira et al. (2005), Pardo (2006)
 - Redes de sumários bons apresentam certas características

Medidas de avaliação intrínseca

- ACL 2005 & RANLP 2005
 - "Graphs everywhere", Dragomir Radev
 - Recuperação de informação, perguntas e respostas, avaliação de sistemas de PLN
- CILing 2006
 - "Random walks on text structures", Rada Mihalcea
 - Sumarização de textos, word sense disambiguation
- HLT/NAACL 2006
 - Workshop: "Textgraphs: Graph-based Algorithms for Natural Language Processing"
- Elegância e simplicidade dos grafos
- Grafos são estruturas muito estudadas
- Em muitos casos, deixa-se de lado a complexidade linguística

Medidas de avaliação intrínseca

- Recentemente...
 - Método da pirâmide (Nenkova e Passonneau, 2004)
 - Codificação prévia de conteúdo
 - *Basic elements* (Hovy et al., 2006)

Avaliação intrínseca

- Necessidade de cópuz
 - Textos e sumários de referência
- Edmundson (1968)
 - Corpus de treino: 200 textos, de ciências físicas, biológicas, humanas e da informação
 - Corpus de teste: 200 relatórios de química, 100-3900 palavras
- Kupiec et al. (1995)
 - 188 documentos científicos/técnicos, aproximadamente 86 sentenças cada
- Teufel e Moens (1997)
 - 202 artigos da *Computational Linguistics*

Avaliação intrínseca

- Necessidade de cópuz
 - Textos e sumários de referência
- Marcu (1997)
 - 5 artigos da *Scientific American*, 161-725 palavras
- Jing et al. (1998)
 - 40 textos jornalísticos (TREC)
- TIPSTER SUMMAC (1998)
 - 183 artigos de Linguística Computacional (ACL)

Avaliação intrínseca

- Necessidade de cópuz
 - Textos e sumários de referência
- Rino et al. (2004)
 - 100 textos jornalísticos
- DUC (2003)
 - Aproximadamente 620 textos jornalísticos

Avaliação extrínseca

- Sumários em contexto
 - Categorização de textos
 - Perguntas e respostas
 - Recuperação de informação
- Os sumários podem não ser bons para o ser humano, mas podem ser bons para a máquina
 - Às vezes, textualidade não é necessária

Avaliação extrínseca

- Categorização de textos (Mani et al., 1998)
 - Atribuir uma classe aos textos: economia, informática, política, etc.
 - Em vez do humano/computador processar o texto todo, processa somente o sumário
 - Taxa de acerto deve aumentar
 - Menos informação irrelevante
 - Tempo demandado deve diminuir
 - Menos informação para processar

Avaliação extrínseca

- Perguntas e respostas (Morris et al., 1992; Hovy e Lin (2000))
 - Preparam-se perguntas para um grupo de textos
 - Humanos respondem as perguntas
 - Sem ler nada
 - Lendo os sumários
 - Lendo os textos

Avaliação extrínseca

- Recuperação de informação (Mani et al., 1998; Tombros e Sanderson, 1998; Jing et al., 1998)
 - Duas abordagens
 - Indexação de sumários em vez dos textos completos
 - Aumento da taxa de acerto na recuperação
 - Somente informação relevante é indexada
 - Apresentação de sumários junto com os resultados da busca
 - Aumento da satisfação do usuário

Futuro da avaliação

- Avaliação
 - Extrínseca
 - *Off-line*
 - *Black-box*
 - Comparativa
 - Multidocumento
- DUC 2005
 - Difícil superar o método *baseline*
 - "Métodos cada vez mais complicados para se selecionar a primeira sentença dos textos"
 - Estagnação da área?
 - Mudança de paradigma



Futuro da avaliação

- Avaliações conjuntas
 - Avaliam o estado da arte
 - Ditam direções de pesquisa



Sumarização Automática: Métodos e Avaliação



www.nilc.icmc.usp.br/~thiago
tasparado@icmc.usp.br