

ExpressMath Online: Um sistema web para o tratamento de dados de expressões matemáticas manuscritas

Davi Souto Grangeiro

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Mestrado em Ciência da Computação

Orientador: Prof^a. Dr^a. Nina S. T. Hirata

São Paulo, Agosto de 2015

ExpressMath Online: Um sistema web para o tratamento de dados de expressões matemáticas manuscritas

Esta é a versão original da dissertação elaborada pelo candidato Davi Souto Grangeiro, tal como submetida à Comissão Julgadora.

Agradecimentos

Agradeço primeiramente a Deus por todos os passos e todas as conquistas de minha vida.

À minha família, por todos os valores e incentivo no decorrer dos anos.

À Priscila Wiltemburg, por todo carinho, incentivo e força que serviram de motivação no desenvolvimento deste trabalho.

À Prof^a Dr^a Nina S. T. Hirata pelos ensinamentos, auxílio, paciência e motivação, conjunto de elementos essenciais para minha evolução. Agradeço pela chance a mim concedida para a realização deste e pela excelente orientação e condução ao decorrer do meu mestrado.

A Frank Dennis Julca Aguilar pela ajuda, esclarecimentos e contribuições, fundamentais no desenvolvimento da minha pesquisa.

A todos amigos e companheiros que contribuíram e me incentivaram direta e indiretamente na luta por esta conquista.

A todos, muito obrigado.

Resumo

GRANGEIRO, Davi S. **ExpressMath Online: um sistema web para o tratamento de dados de expressões matemáticas manuscritas**, 2015. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2015.

O problema de reconhecimento de expressões matemáticas manuscritas tem atraído o interesse de muitos pesquisadores recentemente. Dado esse aumento, é natural que diversos sistemas de reconhecimento surjam junto com os trabalhos de pesquisa que são realizados. Uma consequência interessante disso é o fato de muito tempo ser despendido na construção repetida de aplicações base para o tratamento dos dados e funcionalidades primárias relacionadas ao reconhecimento de expressões, tais como coleta, persistência, gestão de resultados e avaliação dos métodos. Um sistema *online* para o tratamento desse tipo de dados agilizaria as pesquisas nessa área, uma vez que disponibilizaria a base para o desenvolvimento e avaliação de novos métodos e ao mesmo tempo evitaria a geração de diferentes padrões que dificultam comparações.

Além disso, para a análise e avaliação dos métodos e técnicas de reconhecimento de expressões matemáticas, é necessário um conjunto representativo de amostras de expressões anotadas com informações relevantes para a avaliação do reconhecimento. Atualmente, a criação de bases de dados de tamanho expressivo é considerado um desafio para os avanços na área de reconhecimento de expressões matemáticas manuscritas.

Neste trabalho propomos o ExpressMath Online, um sistema para o tratamento de dados de expressões matemáticas manuscritas *online*. O sistema é composto por uma série de funcionalidades básicas para o tratamento dos dados de expressões e avaliação dos resultados de métodos de reconhecimento, que podem ser adicionados ao sistema de forma descomplicada. Além disso, propomos um método baseado no casamento de expressões para facilitar a anotação de múltiplas instâncias de uma mesma expressão. A ideia do casamento consiste em associar os traços de uma expressão transcrita ao correspondente símbolo na respectiva expressão modelo. Uma vez realizado o casamento, todas as informações anotadas na expressão modelo podem ser automaticamente transferidas para a expressão transcrita. O método de casamento foi integrado ao ExpressMath Online e foi utilizado para testar as funcionalidades do sistema.

Abstract

GRANGEIRO, Davi S. **ExpressMath Online: a web system to handle handwritten mathematical expressions data**. 2015. Dissertation (Master Program) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2015.

The problem of handwritten mathematical expression recognition has attracted the interest of many researchers recently. Given this increase, it is natural that several recognition systems arise along with the research works that are done. An interesting result is that a lot of time is spent on repetitively building the application basis for data processing and primary functionality related to the recognition of expressions, such as collection, persistence, management of resulting data and evaluation methods. An online system for the treatment of such data would speed up research in this area since it would provide the basis for the development and evaluation of new methods and at the same time it would prevent the creation of different standards that make comparisons difficult.

Moreover, for the analysis and evaluation of mathematical expression recognition methods and techniques, a representative sample set of annotated expressions with information relevant to the recognition evaluation is necessary. Currently, the creation of databases with significant size is considered a challenge to the advances in the field of handwritten mathematical expression recognition.

In this work we propose ExpressMath Online, a system for the treatment of online handwritten mathematical expression data. The system consists of a series of basic functions for the processing of expression data and for the evaluation of results from recognition methods, which can be added to the system in a simple way. In addition, we propose a method based on expression matching to facilitate the annotation of multiple instances of an expression. The matching idea consists in associating the strokes of a transcribed expression to the corresponding symbol in the corresponding model expression. Once the matching is computed, all the ground-truth data in the model expression can be automatically transferred to the transcribed expression. The matching method has been integrated into ExpressMath Online and it was used to test the system features.

Sumário

Lista de Figuras	xi
Lista de Tabelas	xv
1 Introdução	1
1.1 Motivação	3
1.2 Objetivos	5
1.3 Contribuições	5
1.4 Organização do Trabalho	6
2 O processo de criação de EMs anotados com ground-truth	7
2.1 Expressões matemáticas online	7
2.1.1 Reconhecimento de EMs	7
2.1.2 Avaliação de sistemas de reconhecimento de EMs	8
2.1.3 Anotação de gabarito (<i>ground-truth</i>)	10
2.2 Casamento de expressões proposto em Honda (2013)	10
2.2.1 Formulação como problema de associação	11
2.2.2 Pré-processamentos de EMs	11
2.2.3 Similaridade de formas - Shape context	12
2.2.4 Cálculo dos custos de casamento	14
2.2.5 Resultados obtidos	14
3 Sistema para tratamento de dados de expressões matemáticas	17

3.1	ExpressMath Online, um sistema para tratamento de dados de expressões matemáticas	17
3.2	Características do Sistema	17
3.3	Modelagem dos dados	19
3.3.1	Domínio funcional (Sistêmico)	20
3.3.2	Domínio de dados de EMs	20
3.3.3	Gestão de resultados	21
3.3.4	Parametrizações	21
3.3.5	Metadados	21
3.4	Fluxo de vida das EMs	22
3.5	Funcionalidades que integram o ExpressMath Online	23
3.6	Usuários do sistema	26
4	Método para casamento de expressões	29
4.1	Pré-processamentos de EMs	30
4.1.1	Representação de expressões por grafos	30
4.1.2	Normalização espacial e de escala	30
4.1.3	Normalização dos pontos do contorno do símbolo	31
4.2	Casamento de expressões sem a suposição de segmentação prévia	32
4.2.1	Etapa 1: Seleção de traços candidatos	32
4.2.2	Etapa 2: Cálculo de verosimilhança	34
4.2.3	Etapa 3: Associação preliminar	35
4.2.4	Etapa 4: Remoção de associações sobrepostas	36
4.2.5	Etapa 5: Processamento de símbolos e traços órfãos	36
4.3	Resultados experimentais	37
5	ExpressMath Online na prática	39
5.1	Criação dos modelos	39
5.2	Transcrição das expressões	40

5.3	Teste de métodos de casamento	41
5.4	Parametrizações	42
5.5	Avaliação dos resultados	45
5.5.1	Erros de casamento mais comuns	47
5.6	Tecnologias utilizadas	51
6	Conclusões	55
6.0.1	Trabalhos futuros	56
	Referências Bibliográficas	57

Lista de Figuras

2.1	Uma árvore de mapeamento para a expressão $\cos \theta = \frac{x}{\sqrt{x^2+y^2}}$	9
2.2	Grade log-polar com 5 <i>bins</i> para o logaritmo da distância e 12 <i>bins</i> para o ângulo θ	13
2.3	Associação correta média global por casamento variando o α da equação 2.5 ente 0.0 e 1.0 (gráfico extraído de Honda (2013)).	15
3.1	Modelo de dados do sistema ExpressMath Online . Nele são apresentados quatro domínios de dados, cada um representado por uma cor dentro do diagrama, onde: lilás para o domínio funcional, cinza para dados de EMs, verde para gestão de resultados e laranja para parametrizações.	19
3.2	Hierarquia de dados de uma expressão matemática dentro do ExpressMath Online , constituída por modelos, expressões, símbolos, traços e pontos, que são as entidades necessárias para se representar de maneira completa todos os dados de uma EM.	20
3.3	Fluxo de vida de uma EM dentro do ExpressMath Online	22
4.1	A figura apresenta a entrada e saída (a e b respectivamente) para o processo de normalização dos pontos do contorno do símbolo para a expressão $n! - 1$	31
4.2	Ilustração do processo de seleção de traços. Em <i>a</i> , apresentamos a posição relativa ao símbolo x da expressão $\cos \theta = \frac{x}{\sqrt{x^2+y^2}}$; em <i>b</i> apresentamos o centróide de x sobreposto na expressão transcrita, juntamente com disco definido pela distância considerada para a construção do conjunto de traços que podem corresponder a x	33
4.3	No gráfico, são apresentados as taxas de acerto de associação de símbolos separados por expressão modelo.	38
5.1	Painel de inserção de expressões modelo.	39

5.2	Funcionalidade de transcrição das expressões modelo. O primeiro painel, localizado na parte superior esquerda da figura é onde são apresentadas as expressões modelo, e no segundo painel é onde as transcrições são coletadas.	40
5.3	Funcionalidade de transcrição das expressões modelo, após a transcrição ter sido coletada.	41
5.4	Painel de teste de casamento entre expressões modelo e transcrições. Nesse painel é possível observar o estado da expressão, além do usuário que transcreveu a expressão.	42
5.5	Painel de teste de casamento de expressões após o casamento ter sido realizado. Neste ponto é possível visualizar como o casamento ocorreu, além da taxa de acerto do casamento.	43
5.6	Painel de parametrizações do sistema. Nele é possível definir os valores dos parâmetros utilizados nos métodos de casamento/reconhecimento.	43
5.7	Opção de adição de um novo parâmetro de usuário dentro do ExpressMath Online. Nela o usuário escolhe o parâmetro que deseja definir além de seu valor.	44
5.8	Opção de adição de um novo parâmetro padrão dentro do ExpressMath Online. Nela o usuário define o nome, descrição, tipo e valor padrão do novo parâmetro.	45
5.9	Painel de configuração das avaliações das requisições de casamento. Dentro deste painel é possível configurar o tipo do gráfico desejado, o conjunto de dados no qual será realizada a análise, além das séries disponíveis que irão compor o gráfico.	46
5.10	Configuração do parâmetro α , definido como 0,75.	46
5.11	Configuração do parâmetro α , definido como 0,9.	47
5.12	Painel de requisições de teste do ExpressMath Online. Vale notar que dentre as requisições inseridas, existe um conjunto delas que é destinado a análise do parâmetro α	47
5.13	Painel onde os gráficos de análise de resultados são apresentados. Esse gráfico é dinâmico, e suas informações são provenientes das configurações realizadas no painel apresentado na figura 5.9. No gráfico que há na figura, podemos observar a análise de duas requisições de casamento que tem a finalidade de observar o impacto da alteração do parâmetro α	48
5.14	Caso de casamento de expressões onde ocorreu uma associação invertida entre duas letras 'c'.	49

5.15	Caso onde o espaçamento entre os símbolos é desigual. Note que no decorrer da expressão, o espaçamento diminui, o que causa um erro no casamento do símbolo 's'.	50
5.16	Três maneiras distintas (a), (b) e (c) de escrever o símbolo '2'.	51
5.17	Na figura (a) apresentamos a transcrição de uma expressão modelo, e na figura (b) apresentamos o resultado do casamento entre as expressões; o erro de casamento ocorre pois a forma de escrita dos símbolos difere entre as expressões modelo e transcrição.	52
5.18	Caso de erro de casamento onde restaram traços e símbolos órfãos.	53
5.19	Erro de casamento devido a semelhança entre o subgrupo do sinal de adição (-) com o símbolo fracionário.	54

Lista de Tabelas

- 1.1 Resultados da competição *ICDAR 2013 CROHME: Competition on Recognition of Online Handwritten Mathematical Expressions* (tabela extraída de Mouchère *et al.* (2013)). 2

- 3.1 Matriz de acesso às funcionalidades por perfil. Na tabela, os perfis são descritos por números, onde respectivamente significam: Voluntários, Gerenciadores de EMs, Desenvolvedores de métodos de rotulação e Administradores. 28

Capítulo 1

Introdução

Com o advento de tecnologias de interface de entrada de dados *touch-screen*, houve um grande aumento de pesquisas no desenvolvimento de técnicas para reconhecimento de escrita. Como este tipo de dispositivo permite entre outras, esse tipo de entrada de dados, existe uma grande tendência ao desenvolvimento de aplicações para dispositivos baseados nessa tecnologia.

Existe uma grande diversidade dentre os elementos que compõem a escrita, como texto, diagramas, fórmulas, gráficos, expressões matemáticas, etc. Alguns se mostram mais adaptáveis às tecnologias convencionais de entrada de dados, como teclados e mouses. Porém existem elementos de escrita, tais como as expressões matemáticas (EMs), que necessitam de conhecimento de alguma linguagem específica, tal como \LaTeX , para serem descritas. Uma alternativa interessante seria possibilitar a entrada de dados via escrita manual, que é considerado uma maneira mais simples e natural para descrever expressões matemáticas.

Atualmente, diversos trabalhos se dedicam ao estudo do reconhecimento da escrita. Dentre esses trabalhos um tema de destaque é o reconhecimento de EMs manuscritas. Esse tema tem grande importância em diversas situações, tais como prover uma interface de inserção mais natural para sistemas de resolução e tipografiação de expressões matemáticas; leitura automática de documentos para pessoas portadoras de deficiência e reconhecimento de textos científicos, que geralmente incluem expressões matemáticas.

Podemos dividir o campo de reconhecimento de EMs em três grandes tópicos de estudo: segmentação dos símbolos, reconhecimento dos símbolos e análise estrutural. Dado um conjunto de traços que corresponde a uma EM manuscrita, a segmentação consiste na transformação que leva o conjunto de traços a um conjunto de símbolos; o reconhecimento do símbolo trata da atribuição de uma classe ao símbolo segmentado, associando assim uma identidade ao símbolo; a análise estrutural estuda as relações entre os símbolos pertencentes à expressão, informação utilizada para extrair a semântica da EM.

Diversos trabalhos vem sendo publicados desde a década de 90 (Awal *et al.*, 2009; Chan e Yeung, 2000; Hirata e Julca-Aguilar, 2015; Huang e Kechadi, 2007;

Koschinski *et al.*, 1995; Matsakis, 1999; Toumit *et al.*, 1999; Xie, 2007), e apesar de muitos avanços alcançados, a taxa de reconhecimento ainda não atingiu patamares satisfatórios. Tal fato pode ser evidenciado, por exemplo, analisando-se os resultados divulgados na competição *ICDAR 2013 CROHME: Competition on Recognition of Online Handwritten Mathematical Expressions* (Mouchère *et al.* (2013)) e reproduzidos na tabela 1.1.

TABLE IV. EXPRESSION-LEVEL EVALUATION, SHOWING $\leq 0..3$ SYMBOL OR MATHML STRUCTURE LABEL ERRORS (671 EXPRESSIONS)

	Correct (%)	≤ 1 error	≤ 2 errors	≤ 3 errors
VII	60.36	80.33	84.95	86.14
IV	23.40	37.85	44.71	47.84
II	19.97	34.13	40.83	42.92
III	9.39	18.48	24.14	27.27
VI	8.35	19.08	24.44	26.23
I	2.68	9.69	16.24	20.72
VIII	18.33	32.04	40.24	42.92
V	14.31	24.74	32.19	36.21

Tabela 1.1: Resultados da competição *ICDAR 2013 CROHME: Competition on Recognition of Online Handwritten Mathematical Expressions* (tabela extraída de Mouchère *et al.* (2013)).

Na competição acima mencionada, diversos sistemas de reconhecimento de EMs foram testados. A maior taxa de acerto em reconhecimento EMs, de 60.36%, foi obtido por um sistema que utilizou dados externos aos fornecidos pela organização da competição. Entre os sistemas que utilizaram apenas os dados fornecidos, a maior taxa de acerto obtida foi de apenas 23.40%. Esses resultados estão longe do buscado por sistemas de reconhecimento, o que motiva pesquisas nesse tema.

Além do problema ser difícil devido à grande quantidade de símbolos que precisam ser reconhecidos, à complexidade do arranjo bidimensional da estrutura e à variação intrínseca da escrita, outra dificuldade que tem sido apontada como uma das grandes razões para o lento avanço nessa área é a falta de base de dados públicos de EM's. Tal falta dificulta uma avaliação completa de um sistema e impede comparações e reprodução de diferentes abordagens.

Essa falta vem sendo gradativamente sanada por meio de esforços de alguns grupos de pesquisa. Atualmente podem ser encontradas bases públicas como por exemplo **ExpressMatch**: Julca-Aguilar e Hirata. (2012), **MfrDB**: Bresler *et al.* (2012) e **MathBrush**: Labahn *et al.* (2008). O conjunto de dados utilizados na competição *CROHME 2013* é, de fato, construído pela junção de partes dessas bases e constitui um dos primeiros esforços de se disponibilizar uma base padronizada.

Apesar disso, para uma efetiva avaliação dos reconhecedores é necessário uma base muito maior que as existentes atualmente. Cabe ressaltar que as expressões contidas nessas bases são em sua grande maioria limitadas quanto à variedade de símbolos, tamanho (número de símbolos) e estruturas (por exemplo, matrizes em geral não fazem parte dessas bases).

Contudo, a criação de base de dados para treinamento e teste de reconhecedores não é tarefa trivial. Além do processo de coleta dos manuscritos que é trabalhosa por envolver a necessidade de colaboração de voluntários, outra atividade trabalhosa é a associação de *ground-truth* às expressões. Uma vez que existe o interesse geral de se avaliar a capacidade de reconhecimento de determinada abordagem em diferentes níveis de detalhes (segmentação de símbolo, reconhecimento de símbolo, relação espacial entre símbolos, *baselines*, subestruturas e estrutura global, por exemplo), gabaritos correspondentes a cada um desses níveis de detalhe precisam ser associados às expressões. Esta é uma tarefa demorada e cansativa, além de sujeita a erros.

1.1 Motivação

Dado o aumento do interesse na área de reconhecimento de EMs, é natural que diversos sistemas de reconhecimento surjam junto com os trabalhos de pesquisa. Um fato interessante decorrente disso é, o tempo despendido na construção de aplicações base para o tratamento dos dados e funcionalidades primárias relacionadas ao reconhecimento de EMs, tais como coleta, persistência, gestão de resultados e avaliação dos métodos. Assim, um sistema base que disponibilizasse essas funcionalidades seria muito útil e agilizaria as pesquisas nessa área de uma forma geral, abrangendo um grande número de pessoas.

Apesar de existirem algumas iniciativas que visam estabelecer padrões para facilitar a pesquisa em reconhecimento de escrita, tais como o formato InkML (<http://www.w3.org/TR/InkML/>) para escrita online, e as competições CROHME (Competition on Recognition of Online Handwritten Mathematical Expressions), não há hoje, uma iniciativa de se desenvolver um sistema web para o tratamento desse tipo de dados. Tal iniciativa poderia contribuir com avanços na área, pois permitiria que as pesquisas fossem realizadas mais rapidamente uma vez que toda a infraestrutura sistêmica necessária para o tratamento dos dados já estaria pronta. Além disso, o fato de um sistema estar disponível na internet aumentaria significativamente sua visibilidade.

Além disso, conforme mencionado anteriormente, uma das dificuldades apontadas recentemente na área de reconhecimento de EMs é a baixa quantidade de bancos de dados públicos. As bases são geralmente limitadas a características e aspectos específicos.

A literatura mostra que de forma geral, a fim de testar e avaliar técnicas de reconhecimento, os pesquisadores constroem seus próprios bancos de dados. Isso tende a gerar bancos de dados limitados, muitas vezes estabelece um viés do método proposto aos dados, além de dificultar a comparação entre técnicas de reconhecimento.

Em [Honda \(2013\)](#), foi proposto uma abordagem baseada em casamento de expressões para a rotulação de expressões, que foi aperfeiçoada subsequentemente e descrita em [\(Hirata e Julca-Aguilar, 2015\)](#). A ideia básica consiste em se criar expressões modelo

e associar *ground-truth* manualmente. Voluntários são então convidados a transcrever essas expressões modelo. Para rotular as expressões transcritas, elas são casadas símbolo a símbolo com a respectiva expressão modelo. Uma vez que tal casamento é obtido, é possível repassar automaticamente toda informação de todos os níveis de gabarito da expressão modelo para a expressão transcrita.

Desta forma, várias amostras de uma mesma expressão podem ser geradas apenas realizando-se a coleta das mesmas, sem a necessidade de se anotar cada uma delas. Isso abre possibilidade para se avaliar por exemplo a robustez de sistemas de reconhecimento em relação a variações de escrita, ou o reconhecimento restrito a um grupo de usuários ou a um tipo de dispositivo, entre outras possibilidades.

Os resultados obtidos pelo método proposto em Honda (2013) e (Hirata e Julca-Aguilar, 2015) atingem taxa média de associação correta de símbolos superior a 99%. O método está integrado ao software **ExpressMatch**¹ com o qual foi gerada uma base com 957 transcrições de um conjunto de 64 expressões modelo.

Porém, o referido método impõe uma restrição severa: ele pressupõe que os símbolos estão corretamente segmentados não apenas na expressão modelo mas também em suas transcrições (garantindo, portanto, que o número de símbolos na expressão transcrita é exatamente igual ao da expressão modelo).

Durante a coleta das expressões utilizando o sistema **ExpressMatch**, os usuários são induzidos a escreverem de forma que os símbolos sejam corretamente segmentados. Especificamente, o sistema considera que sempre que o intervalo de tempo entre o fim de um traço e início do traço subsequente é maior que um dado limiar, os traços pertencem a símbolos distintos (isto é, neste caso o segundo traço representa o início de um novo símbolo na escrita). Caso o intervalo de tempo seja menor que o limiar, os traços são considerados como sendo de um único símbolo. Além disso, o sistema desenha um retângulo envolvente (*bounding box*) ao redor do conjunto de traços que está sendo considerado como um único símbolo. Desta forma, o usuário pode perceber e realizar eventuais correções imediatamente caso traços sejam agrupados de forma incorreta.

Como esse tipo de restrição nem sempre pode ser imposto, há necessidade de se desenvolver métodos capazes de lidar com expressões nas quais os símbolos não estão individualmente segmentados. Além disso, há necessidade de se facilitar o processo de verificação e confirmação do casamento, para se ter certeza de que as anotações estão corretas. Uma terceira necessidade é avaliar os métodos de anotação automática em si. Uma vez que o volume de dados usados para validar os algoritmos é muito grande, uma avaliação manual é inviável. É desejável a criação de mecanismos que facilitem a análise de desempenho de diferentes soluções (diferentes tipos de algoritmos, ou mesmo ajuste de parâmetros dos algoritmos).

¹code.google.com/p/express-match/

1.2 Objetivos

O principal objetivo deste trabalho é promover avanços para resolver os problemas listados acima e contribuir para o avanço dessa área de reconhecimento de expressões matemáticas online.

Os objetivos específicos são:

1. Desenvolver um sistema para o tratamento de dados de expressões matemáticas, a fim de facilitar o processo de criação de conjuntos de EMs anotados com ground-truth e agilizar o processo de análise de resultados de diferentes abordagens. O processo inclui os seguintes passos principais:
 - registro de metadados de usuário/condições de coleta
 - coleta de amostras de EMs
 - rotulação de EMs coletadas
 - verificação e confirmação de rotulação
 - incorporação de novos métodos de rotulação
 - avaliação de métodos de rotulação

O sistema **ExpressMatch** citado anteriormente implementa partes desse processo, porém não oferece mecanismos para avaliação automática de métodos de anotação.

2. Propor um método para casamento de expressões que não parte do pressuposto de que os símbolos das expressões transcritas, a serem casadas, estão corretamente segmentadas.

Como já mencionado, os métodos desenvolvidos previamente supõe que os símbolos estão corretamente segmentados.

3. validar o sistema desenvolvido, testando o método de casamento desenvolvido.

1.3 Contribuições

Realizamos um estudo dos métodos e sistemas anteriormente propostos relacionados ao problema tratado nesta dissertação.

Desses estudos, e visando os objetivos propostos, resultaram as seguintes contribuições:

- uma descrição do processo de rotulação de EMs e elementos que impactam sua criação e avaliação;

- um modelo de gestão de dados a fim de suprir as necessidades atuais de se tratar EMs de uma maneira estruturada;
- um sistema web denominado neste texto como **ExpressMath Online**, com o modelo de gestão de dados e todas as funcionalidades desde a coleta, produção de gabaritos verificados, e análise de desempenho de métodos de rotulação; e
- um algoritmo para casamento de EMs quando não é pressuposto que os símbolos estão previamente segmentados.

As implementações do sistema descrito neste trabalho, assim como todos os insumos necessários para construir o sistema estão disponíveis em <https://bitbucket.org/davigrangeiro/express-match>.

1.4 Organização do Trabalho

No capítulo 2 apresentaremos uma breve introdução ao processo de criação de gabaritos anotados com ground-truth, além de destacar alguns trabalhos da área. No capítulo 3 apresentaremos o sistema **ExpressMath Online**, que tem como objetivo preencher a lacuna que há hoje em sistemas de reconhecimento de EMs. No capítulo 4 apresentaremos um método alternativo para o casamento das EMs que não impõe nenhuma pré-condição sobre segmentação prévia dos símbolos. No capítulo 5 apresentaremos o sistema **ExpressMath Online** na prática, sua utilização e funcionalidades. Finalmente, no capítulo 6 apresentaremos as conclusões deste trabalho.

Capítulo 2

O processo de criação de EMs anotados com ground-truth

O conjunto das expressões matemáticas manuscritas pode ser dividido em dois subgrupos, *offline* e *online*. O primeiro é obtido pela captura de dados estáticos a partir de técnicas de digitalização. Já o segundo é obtido através de dispositivos que permitem uma entrada de dados natural à escrita, tais como tablets, mesas digitalizadoras, etc. Podemos dizer, em certo sentido, que a escrita *online* se assemelha a uma imagem vetorial enquanto a *offline* se assemelha a uma imagem matricial.

2.1 Expressões matemáticas online

A escrita *online* tem como unidade básica o traço, onde cada traço, na verdade é um conjunto de pontos amostrados durante o processo de escrita. O número de pontos em cada traço e o espaçamento entre pontos depende não apenas da trajetória do traço, mas também da velocidade de escrita e da sensibilidade do hardware usado na coleta.

Definida o que é a escrita *online*, podemos agora definir que uma EM *online* é formada por uma sequência de traços, divididos em subgrupos, que por sua vez representam os símbolos presentes na EM. Os símbolos podem conter traços não necessariamente consecutivos no tempo, uma vez que o usuário pode escrever de forma arbitrária. Por exemplo, na expressão $t = i^2 + i/2$, um usuário pode colocar os pingos nos i 's ou cortar o t após ter finalizado a escrita da EM. De forma geral, essa segregação temporal dos traços de um símbolo ocorre frequentemente na escrita.

O subconjunto de EMs *online* é mais rico em dados quando comparado ao *offline* (uma vez que possibilita a captura de diversos metadados no momento da escrita).

2.1.1 Reconhecimento de EMs

Em geral, o problema de reconhecimento de expressões matemáticas *online* pode ser dividido em 3 subproblemas importantes:

- Segmentação de símbolos

Consiste em particionar um conjunto de traços de forma que cada parte corresponda a um único símbolo da expressão. Assim, traços de um símbolo não podem fazer parte de partes distintas.

- Classificação (identificação) dos símbolos

Consiste em atribuir identidade (classe) aos símbolos já segmentados.

- Análise estrutural

Consiste em analisar as relações entre os símbolos da EM, a fim de extrair informações semânticas da expressão.

Os métodos de reconhecimento mais elaborados utilizam abordagens holísticas, nos quais esses subproblemas são tratados de modo integrado de forma que informações inerentes a um deles ajude os outros e vice-versa.

O objetivo final do reconhecimento é gerar uma árvore de mapeamento da EM, suficiente para descrever a estrutura da equação. A partir dessa árvore, é possível escrever a expressão em outros formatos, como \LaTeX . Um exemplo dessa árvore é apresentado na figura 2.1

Em geral, o reconhecimento de EMs online é mais simples que as de EMs offline pois pode-se explorar informações referentes à dinâmica de escrita (não disponível no caso de escrita *offline*).

Resultados recentes de competições de reconhecimento de EMs online ([Mouchère et al. \(2013\)](#)) mostram que ainda há muito para se avançar até termos um sistema de reconhecimento "usável".

2.1.2 Avaliação de sistemas de reconhecimento de EMs

Além da dificuldade do problema de reconhecimento em si, avaliar sistemas desse tipo também pode ser encarado como um desafio nessa área. A avaliação de desempenho permite mensurar os progressos do desenvolvimento de técnicas e serve como indicador para planejamento de futuras pesquisas no tema. Essas técnicas podem indicar e apontar razões de erros, facilitando assim seu aprimoramento, além de proporcionar uma estimativa do desempenho esperado em aplicações reais.

Segundo [Chan e Yeung \(2000\)](#), as técnicas de avaliação de desempenho de reconhecimento de EMs podem ser agrupadas em três grandes categorias:

- teste do método sobre um conjunto de expressões e categorização dos resultados por expressão;

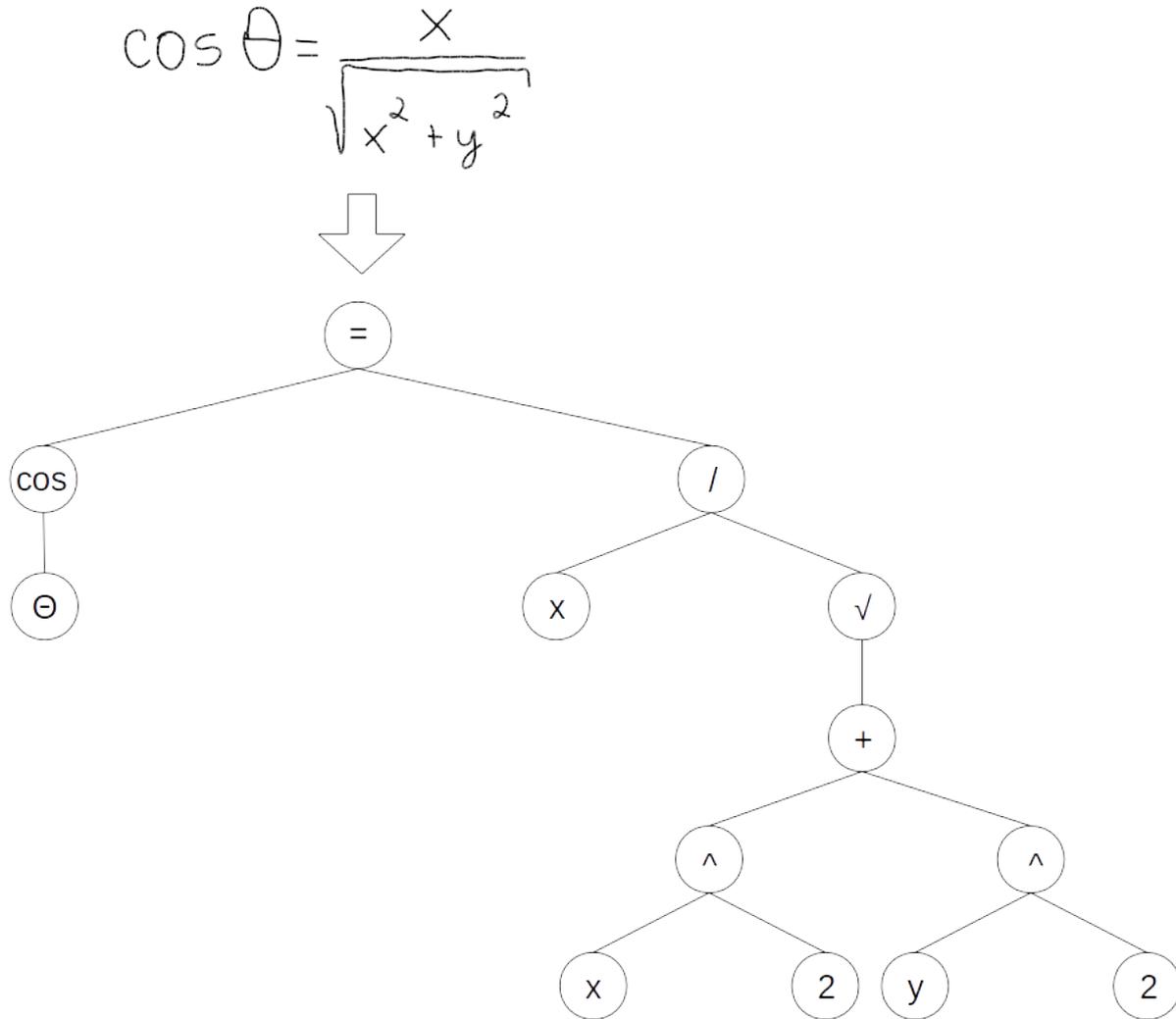


Figura 2.1: Uma árvore de mapeamento para a expressão $\cos \theta = \frac{x}{\sqrt{x^2 + y^2}}$.

- teste do método sobre um conjunto de expressões e observação do desempenho relativo a uma determinada métrica;
- teste em algumas expressões comuns, geralmente escritas por poucos escritores. Como resultado é esperado que todas as expressões sejam corretamente reconhecidas. O objetivo para este tipo de avaliação é confirmar que o método funciona corretamente para expressões mais comuns.

Para mensurar e quantificar avaliações, métricas devem ser escolhidas ou definidas. De acordo com Laponte (2008), métricas de avaliação devem apresentar cinco características:

1. devem refletir o objetivo da técnica;
2. devem ser abrangentes, ou seja, devem considerar todos os objetivos da técnica;

3. devem ser compatíveis com avaliações humanas;
4. devem ser aceitas e reconhecidas pela comunidade científica; e
5. devem ser claras a fim de permitir sua reprodução.

Um levantamento de métricas para avaliar as técnicas para reconhecimento de expressões matemáticas manuscritas é apresentado em [Awal et al. \(2010\)](#). Segundo os autores, cinco métricas são mais utilizadas para avaliação:

- segmentação de símbolos
- reconhecimento de símbolos
- reconhecimento de relações espaciais entre um conjunto de símbolos (sobrescrito, subscrito, acima, abaixo, direita horizontal, etc)
- detecção de subexpressões
- reconhecimento de expressões propriamente

2.1.3 Anotação de gabarito (*ground-truth*)

No contexto de expressões matemáticas, a avaliação de sistemas de reconhecimento apenas no nível de expressão não é suficiente para evidenciar pontos fortes e fracos de técnicas de reconhecimento. Ao invés disso, necessitamos de grandes bases de dados de EMs anotadas com gabarito em diversos níveis (Segmentação de símbolo, classificação de símbolo, relações espaciais entre símbolos, *baselines*, etc), o que nos permitiria fazer uma análise mais detalhada dos métodos propostos.

Dada tal necessidade, o formato InkML¹ é amplamente utilizado como padrão para descrição de gabaritos de EMs. Esse padrão é utilizado no trabalho apresentado na seção 2.2.

2.2 Casamento de expressões proposto em [Honda \(2013\)](#)

O método de casamento de expressões proposto em [Hirata e Honda \(2011\)](#); [Hirata e Julca-Aguilar \(2015\)](#); [Honda \(2013\)](#) consiste de uma abordagem baseada em casamento de expressões para a rotulação de expressões. A ideia básica é criar expressões modelo, associar *ground-truth* manualmente às mesmas, após isso, coletar transcrições e fazer a associação símbolo-a-símbolo entre as duas expressões (expressão modelo que foi anotada manualmente e a transcrição do usuário). Uma vez o casamento tendo sido realizado, copiam-se os rótulos da expressão modelo para os correspondentes elementos na

¹<http://www.w3.org/TR/InkML/>

expressão transcrita. Desta forma, várias amostras de uma mesma expressão podem ser geradas apenas realizando-se a coleta das mesmas, sem a necessidade de se anotar cada uma delas manualmente.

No restante desta seção apresentamos a formulação, procedimentos e resultados obtidos no trabalho de (Honda, 2013).

2.2.1 Formulação como problema de associação

Em Honda (2013), as expressões são representadas por grafos nos quais vértices correspondem aos símbolos e arestas expressam as relações espaciais entre os símbolos.

Como mencionado anteriormente, nesse método o usuário é levado a escrever as expressões de forma que a segmentação dos símbolos esteja correta. Dada essa característica, é garantido que ambas as expressões (modelo e transcrição) possuem o mesmo número de elementos. Com isso, o casamento de expressões pode ser visto como um problema de associação simples, uma vez que ambos os conjuntos possuem a mesma cardinalidade. Além disso, como existe uma correspondência de um para um entre seus elementos, podemos enxergar o problema como sendo o casamento de grafos bipartidos.

Em um problema de associação, custos de associação entre os pares devem ser definidos. No contexto de EMs, uma vez que as mesmas são arranjos bidimensionais de pontos, informação estrutural da expressão (posição relativa entre os símbolos da expressão) e características dos símbolos (forma e traço) são informações relevantes para se estabelecer tal custo.

Para resolver o problema de associação no trabalho citado, foi utilizado o algoritmo húngaro Munkres *et al.* (1957) e Kuhn (1955), que resolve o problema em tempo cúbico.

2.2.2 Pré-processamentos de EMs

São aplicados três pré-processamentos para que o casamento entre as expressões ocorra de forma correta, são eles: normalização de escala, normalização espacial e normalização dos pontos dos traços do símbolo.

Normalização de escala e espacial

Como expressões matemáticas são arranjos bidimensionais, compostos por pontos e traços (segmentos de reta), podem facilmente ser representados por grafos. Com isso, essa representação é utilizada para realizar este pré-processamento.

Os vértices do grafo representam os símbolos e as arestas as possíveis relações entre os símbolos. No trabalho de Honda, é utilizado o grafo completo, onde vértices recebem atributos relativos aos símbolos, incluindo sua coordenada no plano 2D e as arestas recebem

atributos como o seu comprimento e orientação.

Portanto, dados dois grafos similares, um denominado grafo de entrada e outro como grafo modelo², o processo de normalização de escala consiste em rearranjar o grafo de forma que o grafo de entrada seja espacialmente similar ao grafo modelo.

Para realizar tal normalização, consideramos as seguintes etapas:

1. Obtém-se o centróide de ambos os grafos a partir de seus *bounding boxes*.
2. O grafo de entrada é transladado de forma que as coordenadas de seu centróide coincidam com as coordenadas do centróide do grafo modelo.
3. A largura e a altura do *bounding box* do grafo modelo também são alterados, independentemente, de forma que os tamanhos coincidam em ambos os grafos. Neste processo, os vértices são escalados proporcionalmente.

Normalização dos pontos do contorno do símbolo

Dada a grande variação da escrita, o método de captura online tende a gerar pontos redundantes e com espaçamento desigual. Para uniformizar o contorno do símbolo, aplica-se a técnica de pré-processamento descrita em Huang *et al.* (2007). Nessa técnica são aplicados três passos para realizar a normalização dos pontos do contorno dos símbolos; são eles:

1. **Passo 1:** Remoção dos pontos duplicados e 'ganchos' das extremidades dos traços
2. **Passo 2:** São filtrados os ruídos com uma técnica de *smoothing*, e então normalizados.
3. **Passo 3:** As distâncias entre os pontos são equalizadas.

De acordo com Huang *et al.* (2007), essa técnica de pré-processamento aprimora o reconhecimento em 10%.

2.2.3 Similaridade de formas - Shape context

Shape context é um método descritor de formas proposto por Belongie *et al.* (2002). O método é utilizado para mensurar a similaridade entre duas formas.

Dado um conjunto de pontos pertencentes à borda de um objeto, o objetivo do método é associar, à cada ponto do conjunto, informações relativas às posições dos demais pontos no contorno da forma.

Naturalmente esperamos que os pontos do contorno de duas formas similares sejam também similares quanto a esse tipo de informação. Para descrever essa informação utiliza-se

²correspondendo respectivamente às expressões *transcrita* e *modelo*

um histograma devido a grande quantidade de pontos pertencentes ao contorno de formas gerais; essa estrutura é o chamado *shape context*.

O histograma do *shape context* é baseado em regiões denominadas *bins*. Esses *bins* são gerados a partir do espaço log-polar referente a um ponto definido no espaço de coordenadas \mathbb{R}^2 . Tais bins podem ser criados de diversas maneiras, desde que a divisão feita seja uniforme e não haja sobreposição entre os bins.

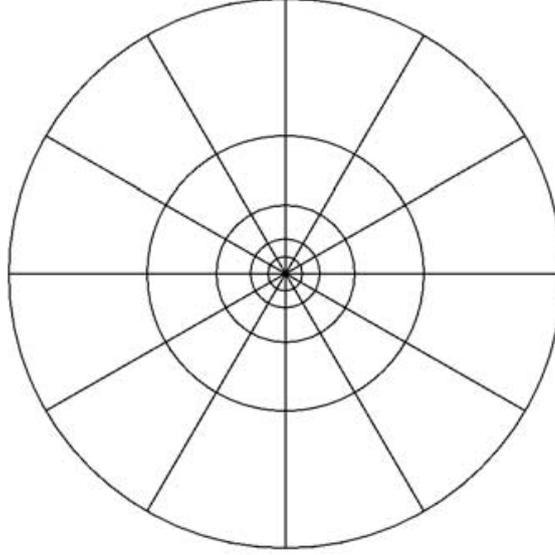


Figura 2.2: Grade log-polar com 5 bins para o logaritmo da distância e 12 bins para o ângulo θ .

Para computarmos os histogramas referentes a cada um dos pontos da forma temos que: Seja p_i um ponto pertencente ao contorno da forma, o histograma a ele relacionado é definido da seguinte forma:

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(k)\} \quad (2.1)$$

onde são contabilizados em cada classe $h_i(k)$ do histograma, os q pontos diferentes de p_i pertencentes ao k -ésimo *bin* do *shape context* referente a p_i .

Belongie et. al definem também uma medida de dissimilaridade entre dois pontos com a prerrogativa de que se duas formas forem bem similares, seus respectivos *shape context* também serão. A medida é baseada em dois pontos p_i e q_j :

$$C_{ij} \equiv C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^K hp(p_i, q_j, k). \quad (2.2)$$

$$hp(p_i, q_j, k) = \begin{cases} \frac{[hp_i(k) - hq_j(k)]^2}{hp_i(k) + hq_j(k)} & \text{caso } hp_i \neq 0 \text{ ou } hq_j \neq 0, \\ 0 & \text{caso contrário.} \end{cases} \quad (2.3)$$

onde K é a quantidade de *bins* do *shape context*, com hp_i e hq_j sendo os histogramas relativos aos pontos p_i e q_j .

Com a medida de dissimilaridade entre 2 histogramas de *shape contexts* definida, pode-se definir a associação dos pontos entre as duas formas da melhor maneira possível, como sendo a que minimiza $H(\pi)$ em 2.4:

$$H(\pi) = \sum_i C(p_i, q_{\pi(i)}) \quad (2.4)$$

no qual π denota uma permutação no conjunto de pontos das formas. A resolução de tal problema é facilmente encontrada aplicando o algoritmo húngaro.

2.2.4 Cálculo dos custos de casamento

Os custos de associação são definidos em termos de deformação entre árvores ARG (Grafos relacionais com atributos). Esses custos são baseados no trabalho de Noma (Noma *et al.*, 2011), no qual é proposto um custo de deformação (Eq. 2.5), que leva em conta informações locais (características do objeto) e informações globais (características estruturais).

Mais detalhadamente, dados dois grafos, G_m (grafo modelo), e G_i (grafo entrada), temos que:

$$c(u_i, v_m) = \alpha c_v(u_i, v_m) + (1 - \alpha) c_e(u_i, v_m) \quad (2.5)$$

na qual $v_m \in V(G_m)$, $u_i \in V(G_i)$ e $\alpha \in \mathfrak{R}$. O parâmetro α ($0 \leq \alpha \leq 1$) é usado para ponderar a importância entre os custos c_v e c_e (de vértice e de aresta, respectivamente) da equação. Esse custo busca expressar a 'deformação' no grafo modelo caso seu vértice v_m seja trocado pelo vértice u_i do grafo de entrada.

Os custos (de vértice e de aresta) são definidos através da diferença entre *shape contexts* (Eq. 2.4). O custo do vértice leva em conta o *shape context* relativo a disposição dos pontos de um símbolo e o custo da aresta leva em conta o *shape context* relativo a disposição dos centróides dos símbolos da expressão.

2.2.5 Resultados obtidos

De acordo com Honda (2013), experimentos realizados em um conjunto composto de 957 expressões é obtida uma taxa média de acerto de 99,7%. Esse resultado foi validado aplicando-se o método em um segundo conjunto de expressões, com 64 amostras, sobre o qual foi obtida uma taxa média de acerto de 99,9%. Esses fatos evidenciam a robustez da técnica proposta. A figura 2.3, extraída de Honda (2013), mostra os resultados do método proposto sobre o conjunto de validação, com a variação do peso das arestas e vértice do grafo (ARG).

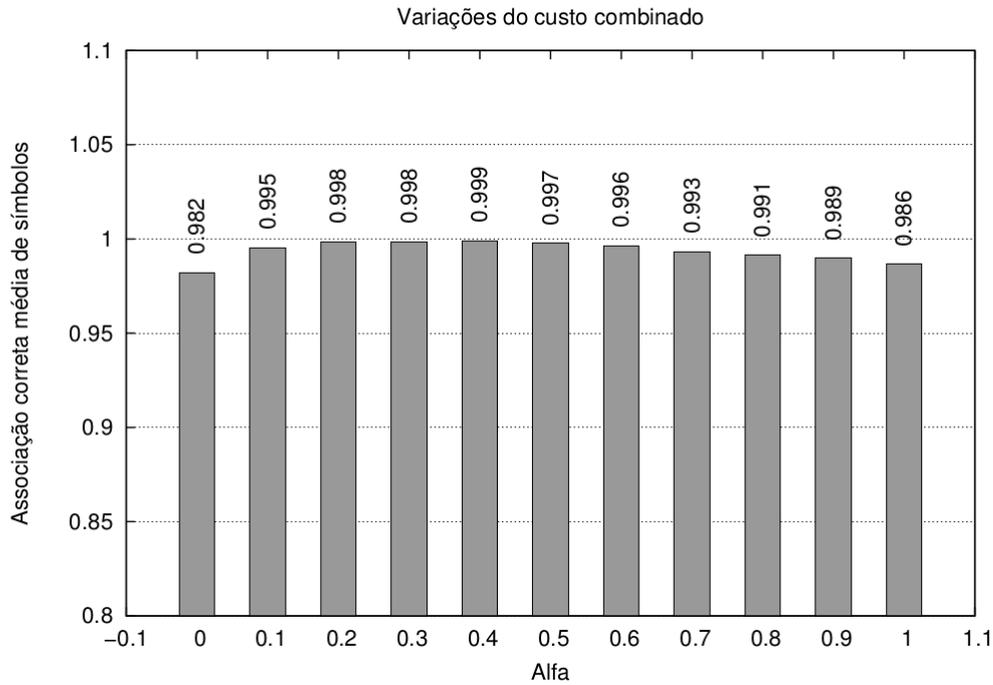


Figura 2.3: Associação correta média global por casamento variando o α da equação 2.5 ente 0.0 e 1.0 (gráfico extraído de Honda (2013)).

O método desenvolvido em Honda (2013) e aperfeiçoado em Hirata e Julca-Aguilar (2015), foi incorporado ao sistema ExpressMatch (Julca-Aguilar e Hirata., 2012), disponível em <http://code.google.com/p/express-match/>. Também encontra-se disponível a base de dados composta por 957 expressões matemáticas anotadas com informações de *ground-truth*.

Capítulo 3

Sistema para tratamento de dados de expressões matemáticas

Ao decorrer deste capítulo, iremos descrever as características e funcionalidades que o sistema **ExpressMath Online** possui e a importância de cada uma delas dentro do processo de tratamento de dados de EMs.

3.1 ExpressMath Online, um sistema para tratamento de dados de expressões matemáticas

Como mencionado anteriormente, um sistema para a gestão e tratamento dos dados de EMs pode colaborar de forma ativa em novas pesquisas nessa área, uma vez que muito tempo é perdido com a construção de sistemas base para o processamento e avaliação de técnicas de reconhecimento de EMs.

O sistema **ExpressMath Online** surge como uma tentativa de agilizar o processo de pesquisa na área de reconhecimento de EMs online. Esse sistema visa disponibilizar todo *background* funcional necessário para o tratamento desse tipo de dados. Dentro desse sistema, estão presentes funcionalidades como coleta de dados, anotação de amostras com *ground-truth* baseada em casamento de expressões, avaliação de métodos de reconhecimento entre outras.

3.2 Características do Sistema

A seguir descreveremos as cinco principais características que o **ExpressMath Online** possui, assim como a motivação e influências destas características dentro do sistema.

- **Sistema Web**

Uma das principais preocupações na construção de sistemas para o tratamento de

dados é a disseminação do seu uso, uma vez que quanto maior a quantidade dos dados, maior é o leque de possibilidades de estudos sobre os dados. Com isso, uma das maiores preocupações na construção do **ExpressMath Online** foi uma ampla utilização pela comunidade. Assim, sua construção é baseada em tecnologias disponíveis para *world wide web*.

- ***Window-based***

Ultimamente, um dos pontos principais na construção de sistemas é a usabilidade. Visando tal característica, a construção da interface gráfica do sistema **ExpressMath Online** foi realizada de maneira intuitiva, de modo que um usuário qualquer conseguisse usar o sistema facilmente. Para isso, ele é baseado em janelas, e a profundidade da navegação dentre as telas é a menor possível. Tais fatos simplificam o entendimento do sistema possibilitando uma melhor experiência de usuário.

- ***Touch-friendly***

Neste trabalho, os objetos de estudo são EMs *online*, com isso, é natural que o usuário necessite de uma interface de entrada de dados *touch-screen*. Dada esta premissa, o sistema **ExpressMath Online** é totalmente *touch-friendly*, portanto, sua interface permite que o usuário comande o sistema utilizando apenas esta entrada de dados.

- **Fácil incorporação de novos dados**

Alguns pesquisadores preferem utilizar sua própria base de dados quando estão analisando dados. Baseado nessa necessidade, o sistema **ExpressMath Online** possibilita a importação de novos dados a partir de um formato amplamente utilizado pela comunidade, o InkML ¹, uma linguagem de descrição de escrita online digital. Assim, é possível importar um conjunto de dados próprio e testar o comportamento das técnicas do sistema nesses dados de uma forma muito simples e intuitiva.

- **Extensibilidade**

Como mencionado anteriormente, a motivação para a construção do sistema é suprir a necessidade de um sistema de tratamento de dados de EM's. Porém, para atender essa necessidade, não basta somente permitir que os usuários incorporem seus dados ao sistema, mas também suas técnicas e métodos de reconhecimento/casamento. Há uma preocupação especial com essa característica, uma vez que essa possibilidade de extensão poderia aumentar de forma considerável a aceitação do sistema como base para o estudo de EMs. Com isso, o sistema **ExpressMath Online** permite, através de uma programação orientada a contratos ² e mecanismos de injeção de dependência, que novas técnicas sejam adicionadas facilmente.

¹InkML: <http://www.w3.org/TR/InkML/>

²Programação orientada em contratos pode ser vista como a utilização de classes abstratas puras a fim de estabelecer os comportamentos esperados dos métodos em uma linguagem orientada a objetos.

3.3 Modelagem dos dados

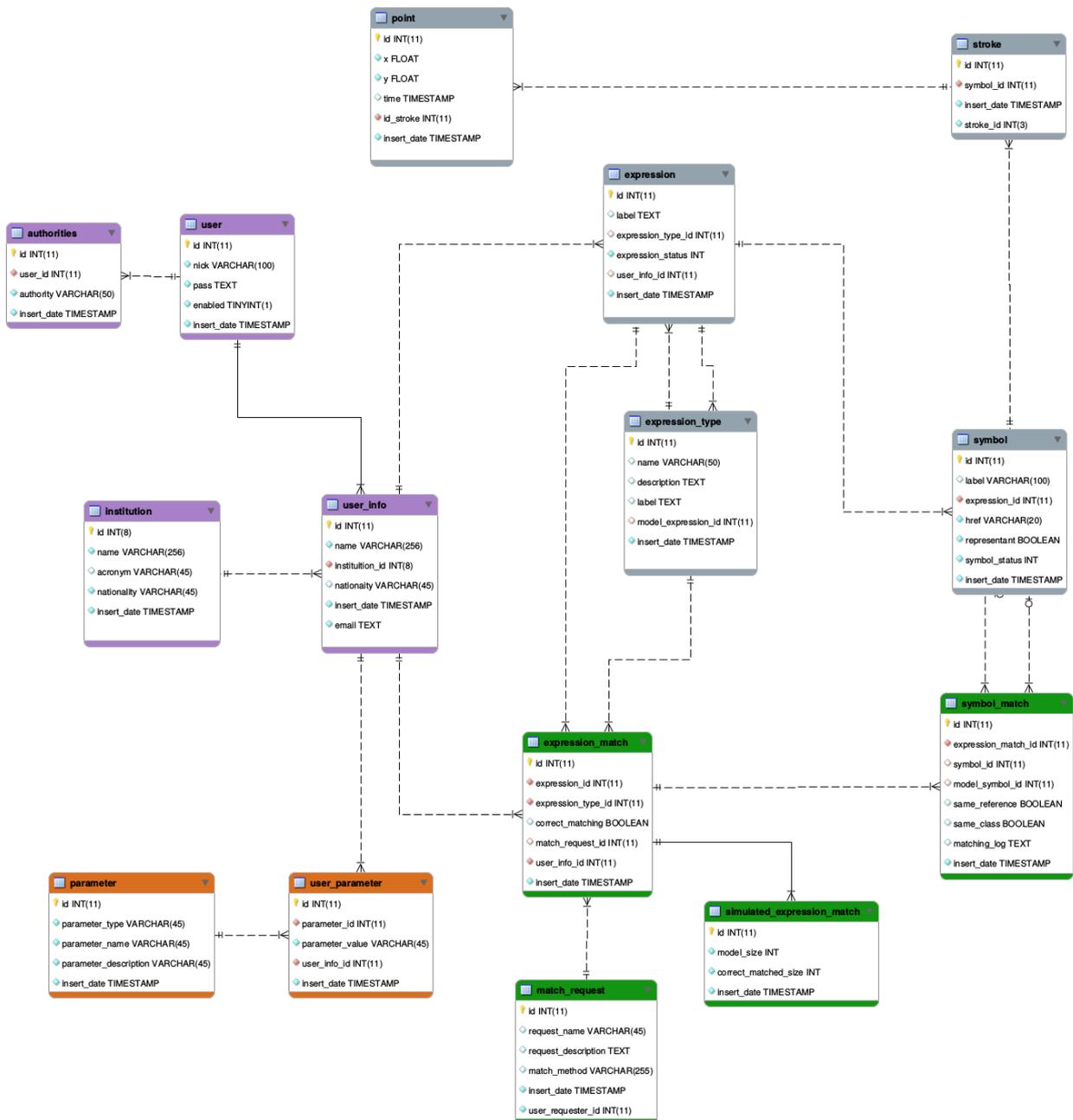


Figura 3.1: Modelo de dados do sistema *ExpressMath Online*. Nele são apresentados quatro domínios de dados, cada um representado por uma cor dentro do diagrama, onde: lilás para o domínio funcional, cinza para dados de EMs, verde para gestão de resultados e laranja para parametrizações.

A modelagem dos dados do sistema ExpressMath Online é subdividida em quatro domínios principais: funcional (sistêmico), dados de EMs, gestão de resultados e parametrizações. Na figura 3.1 é apresentado o modelo completo, onde cada domínio é representado por uma cor, são elas: lilás para o domínio funcional, cinza para dados de EMs, verde para gestão de resultados e laranja para parametrizações. A seguir, descreveremos qual tipo de dado está presente em cada domínio e qual sua função dentro do sistema.

3.3.1 Domínio funcional (Sistêmico)

O domínio funcional dos dados é responsável pelo armazenamento dos dados necessários para o tratamento de *background* do sistema, por exemplo dados de usuários, informações de login, perfis de acesso, entre outros.

3.3.2 Domínio de dados de EMs

Como já dito, neste trabalho o objeto de estudo são EMs *online*. É natural que, para representar de maneira satisfatória este tipo de dado, constituída de traços e pontos, se utilize grafos, onde pontos são representados por vértices e segmentos de reta são representados por arestas.

Tal representação é suficiente para descrever as EMs, porém não possui todas as informações que queremos, como por exemplo quais traços fazem parte de cada símbolo e qual símbolo pertence a qual expressão. Para representar de maneira completa os dados de EMs dentro do **ExpressMath Online**, o domínio de EMs é baseado numa hierarquia de dados apresentada na figura 3.2. Essa hierarquia (baseada em expressões modelo, expressões, símbolos, traços e pontos) visa disponibilizar controle dos dados das expressões e permite, de forma trivial, que dados de *ground-truth* sejam adicionados e extraídos da expressão.

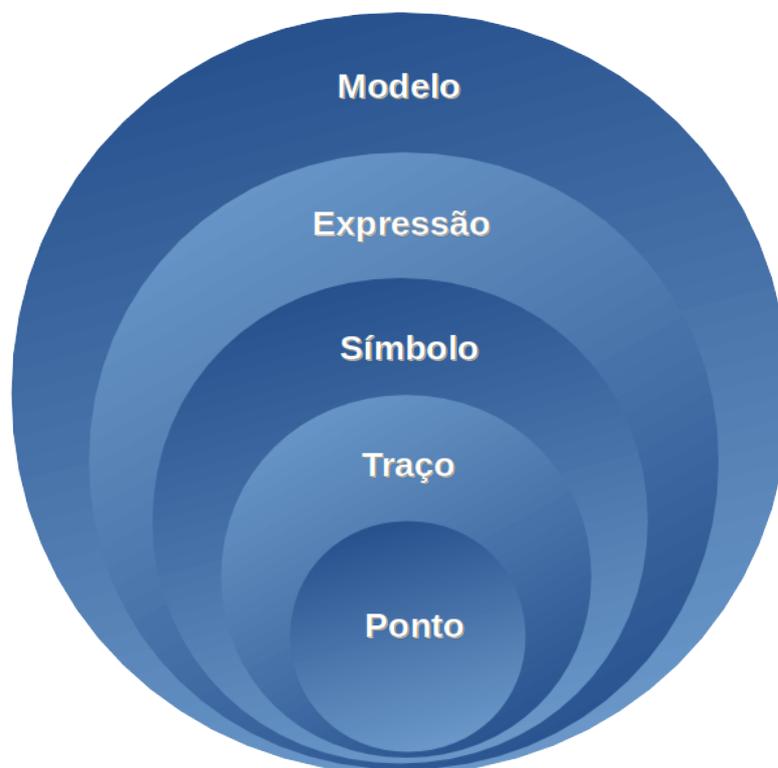


Figura 3.2: Hierarquia de dados de uma expressão matemática dentro do **ExpressMath Online**, constituída por modelos, expressões, símbolos, traços e pontos, que são as entidades necessárias para se representar de maneira completa todos os dados de uma EM.

3.3.3 Gestão de resultados

Devido ao grande volume de dados necessário para avaliação de técnicas de reconhecimento, uma análise manual dos dados se torna inviável. Muitas vezes, pesquisadores extraem resultados para arquivos de texto e os avaliam dentro de algum sistema estatístico, por exemplo a plataforma R. Porém tal abordagem é morosa e cansativa.

Baseado nisso, há, dentro do sistema, um domínio específico para gestão dos resultados. Neste domínio, é possível armazenar informações sobre o reconhecimento das expressões, agilizando e facilitando a extração de resultados e entendimento de falhas/problemas nos métodos ora desenvolvidos.

3.3.4 Parametrizações

Diversas técnicas de reconhecimento necessitam de parametrizações em sua avaliação. Tais parâmetros são geralmente fixados no código fonte dos métodos e avaliados, porém se os parâmetros precisarem ser alterados para alguma avaliação, todo sistema precisa ser contruído novamente. Tal fato é, de forma geral, um problema para os pesquisadores.

Para solucionar este problema, o sistema disponibiliza um domínio específico para parametrizações. Esse domínio permite que parâmetros sejam adicionados/alterados dinamicamente em tempo de execução, possibilitando que diferentes parametrizações sejam avaliadas para determinado método sem a necessidade de uma contrução do sistema como um todo.

3.3.5 Metadados

O domínio de gestão de resultados é capaz de armazenar resultados e possibilita a análise do comportamento de técnicas de reconhecimento/casamento sobre um determinado conjunto de dados. Contudo, podem existir outras características que afetam o reconhecimento, que não são expressas nos dados. O **ExpressMath Online** tenta capturar essas características através da identificação de metadados no decorrer da transcrição da EM.

Os metadados identificados e que fazem parte do **ExpressMath Online** são: a nacionalidade do redator da expressão, uma vez que a grafia pode diferir entre regiões, e a resolução do hardware utilizado na transcrição da EM, pois a diferença entre resoluções pode interferir diretamente na amostragem dos pontos dos traços que compõem a EM.

A ideia é que, ao decorrer do tempo, outros metadados sejam identificados e integrados ao sistema para que os pesquisadores possam analisar seu impacto no reconhecimento das expressões.

3.4 Fluxo de vida das EMs

Como decrevemos anteriormente, a criação de dados anotados com *ground-truth* é realizada, no **ExpressMath Online**, por meio do casamento entre duas expressões, modelo e transcrição. Porém, esta técnica não assegura que todos os símbolos da transcrição foram associados com seus correspondetes na expressão modelo corretamente, uma vez que o casamento é suscetível a erros. Portanto é necessária a validação feita por um usuário para garantir que a expressão foi corretamente anotada.

Pela existência desta característica no sistema, é natural que se estabeleça um fluxo de vida (estados) para a expressão dentro do **ExpressMath Online**. Para detalhar melhor este fluxo, é apresentado na figura 3.3, um diagrama do fluxo dos estados da EM dentro do sistema, assim como as transições entre eles. A seguir, temos as descrição dos estados apresentados no diagrama e como eles são alcançados dentro do sistema.

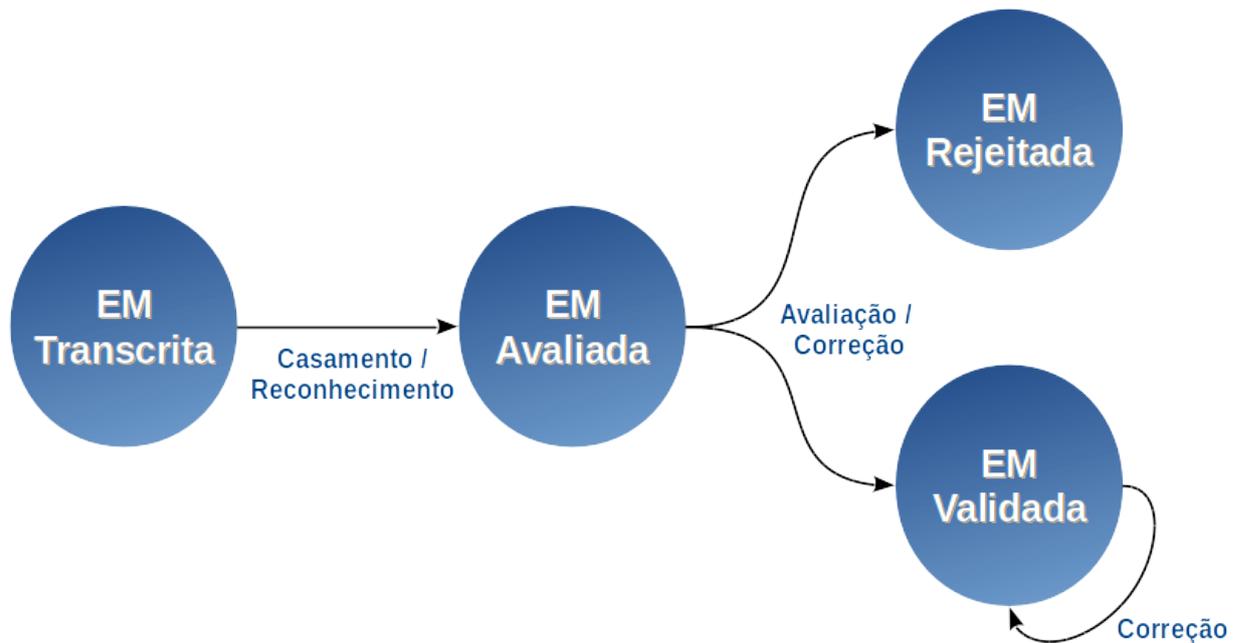


Figura 3.3: Fluxo de vida de uma EM dentro do **ExpressMath Online**.

- **EM Transcrita**

É o estado inicial de uma EM dentro do sistema, ocorre quando um usuário transcreve uma expressão nova ou baseada em um modelo.

- **EM Avaliada**

Após a entrada de uma nova expressão dentro do sistema, ela passa pelo processo de reconhecimento/casamento para a associação de *ground-truth*, após este processo a expressão tem seu estado alterado para EM avaliada.

- **EM Rejeitada**

Como descrito acima, após a avaliação dentro do sistema, as expressões devem passar por um fluxo de avaliação para garantir que os dados foram reconhecidos/casados corretamente. Caso o resultado seja negativo, uma cópia da expressão casada/reconhecida com erros é salva como rejeitada enquanto a expressão original é passada para o processo de correção.

Nesse processo, o usuário apenas confirma se o casamento foi executado perfeitamente ou não. Assim, caso existam erros, a expressão pode seguir seu fluxo de vida e caso contrário seu fluxo de vida termina e então ela passa a incorporar o banco de expressões do sistema.

É salva uma cópia da expressão que foi reconhecida/casada com erro, pois é interessante que o usuário que requisitou o casamento/reconhecimento possa saber qual erro ocorreu. Além disso, com o crescimento da base de dados de expressões rejeitadas, será possível realizar uma análise para encontrar quais são os pontos falhos nas técnicas de uma forma geral, não apenas analisando expressão a expressão.

- **EM Validada**

No processo de validação do reconhecimento/casamento das expressões, podem acontecer dois casos. O primeiro é aquele onde a expressão é corretamente reconhecida/casada e o segundo é aquele onde ocorreram erros no processo.

No primeiro caso, a EM tem seu estado alterado para validada, pois é garantido que não existem erros de reconhecimento/casamento. Já no segundo caso, a EM necessita de correções no casamento/reconhecimento para garantir a consistência dos dados no sistema; portanto, somente após essa correção a expressão tem seu estado alterado para validada.

As expressões validadas passam então a integrar o banco de dados de expressões válidas do sistema, uma vez que é garantido que as informações de *ground-truth* estão corretas nas expressões. Porém, existe a chance de acontecerem erros no processo de correção ou validação da expressão, por exemplo quando o usuário validador não percebe um erro no momento da validação ou comete algum erro no processo de correção, que é muito semelhante a uma anotação manual, sendo então suscetível a erros. Portanto pode ocorrer outro ciclo de validação/correção sobre uma EM que já se encontra em estado validado.

3.5 Funcionalidades que integram o ExpressMath Online

Como mencionamos acima, o sistema é formado por uma série de funcionalidades que visam a agilização do processo de tratamento de dados de EMs. Ao decorrer desta seção, iremos apresentar as funcionalidades presentes no sistema assim como a motivação de sua

inclusão. Além disso, no capítulo 5 apresentaremos o sistema ExpressMath Online na prática, visando ilustrar o processo de coleta dos dados de expressões matemáticas, desde a criação dos modelos até a avaliação dos resultados dos casamentos.

- **Coleta de amostras de EMs baseadas em modelo**

Consiste no processo de coleta de amostras baseada em modelos. Nela, o voluntário transcreve uma expressão modelo informada para que após o processo de casamento, a informação de *ground-truth* seja transferida da expressão modelo para a transcrição, que após o processo de validação/correção é integrada a base de dados do ExpressMath Online.

- **Coleta de EMs**

Consiste no painel de coleta de EMs, no qual o usuário é livre para escrever qualquer tipo de expressão, sem limitação alguma. Após a coleta, a EM será enviada para um processo de reconhecimento e rotulação.

O processo de reconhecimento não é alvo deste trabalho, portanto há dentro do sistema uma interface que permite que métodos de reconhecimento sejam adicionados de uma maneira simples. Assim, pesquisadores que trabalham com reconhecimento de EMs podem utilizar o sistema para testarem suas próprias técnicas.

- **Parametrizações**

É a funcionalidade que permite a adição de parâmetros dentro do **ExpressMath Online**. Nela, o usuário pode escolher uma série de tipos de parâmetros e adicionar novos parâmetros sem a necessidade de alteração alguma no código fonte.

Junto à funcionalidade, é disponibilizado um serviço para busca dos parâmetros do sistema. Assim o desenvolvedor pode se basear neste serviço para utilizar o domínio de parâmetros do sistema na implementação de seus métodos.

- **Processo de casamento de amostras de EMs baseadas em modelos**

Periodicamente, as transcrições das expressões modelo são submetidas ao processo de casamento de EMs com a finalidade de transferir a informação de *ground-truth* da expressão modelo para a transcrição. Nesta funcionalidade, todas as EMs que estiverem com o estado transcrita são submetidas ao processo de casamento, e após o casamento seu estado é alterado para avaliada.

- **Criação de modelos de EMs**

Nesta funcionalidade, os usuários têm a possibilidade de criar novas expressões modelo para integrarem o banco de dados de EMs do sistema. Para isso, o usuário deve inserir uma EM e rotular manualmente seus símbolos e após a inserção e validação dos dados,

a EM passa a ser considerada como um novo modelo para a coleta de expressões baseadas em modelos.

- **Teste de métodos de reconhecimento/casamento (*Offline*)**

Consiste na funcionalidade de submissão de uma requisição de teste sobre todo o conjunto de dados do sistema. Para submeter esta requisição, o usuário precisa informar a motivação para o teste, além do método de casamento/reconhecimento que deseja testar. Os resultados do casamento/reconhecimento são armazenados em banco de dados, dentro do domínio de gestão de resultados, para possibilitar ao usuário realizar a análise dos resultados após a conclusão do processo.

Note que, neste contexto, a palavra *offline* significa que o sistema realizará o casamento/reconhecimento em segundo plano, portanto o usuário deverá analisar os resultados após a conclusão da requisição.

- **Teste de métodos de reconhecimento/casamento (*Online*)**

Consiste no painel de teste de métodos de reconhecimento/casamento. Nele, o usuário pode testar quaisquer métodos que estejam disponíveis dentro do sistema, porém o teste acontece apenas para uma expressão, além de não afetar o estado da expressão. Assim, caso uma nova expressão tenha sido inserida no sistema e ainda não tenha passado pelo processo de casamento/reconhecimento, será possível verificar o comportamento de determinado método sobre essa nova instância. Além disso, também é possível testar qual o impacto de uma alteração de algum parâmetro sobre uma determinada EM.

Note que, neste contexto, a palavra *online* significa que o sistema realizará o casamento/reconhecimento instantaneamente, disponibilizando os resultados ao usuário após seu término.

- **Avaliação de expressões**

No processo de avaliação de expressões, é disponibilizado ao usuário uma lista das expressões que passaram pelo processo de reconhecimento/casamento e precisam ser avaliadas. Após a seleção de uma expressão da lista, é apresentado o resultado do reconhecimento/casamento que foi realizado. Assim, caso não existam erros, o usuário valida a EM e então ela passa a ser parte do banco de dados de expressões válidas do sistema. Caso contrário, uma cópia da EM é salva como rejeitada e a expressão original passa para o processo de correção de casamento/reconhecimento.

- **Correção de expressões**

Quando ocorrem erros no processo de casamento/reconhecimento das expressões, para que a EM seja integrada ao banco de dados, é necessário que aconteça um processo de correção para corrigir os erros que ocorreram no casamento/reconhecimento. Esse

processo é bem parecido com uma rotulação manual, com a diferença que uma parte deles já estarão reconhecidos. Após esse processo o estado da EM é alterado para validada e ela passa a integrar o banco de dados de expressões válidas do sistema.

- **Avaliação de reconhecimento/casamento de EMs**

Como mencionado anteriormente, para agilizar o processo de análise dos resultados de reconhecimento/casamento, existe dentro do **ExpressMath Online**, uma funcionalidade voltada para a análise dos resultados. Nela o usuário pode comparar execuções de diferentes métodos ou de alterações de parâmetros para um determinado método, ou até escolher um determinado conjunto a ser analisado.

- **Processo de adição de novos métodos**

Esse processo visa disponibilizar uma maneira para que os usuários integrem seus próprios métodos dentro do **ExpressMath Online**. Para isso, é disponibilizada ao usuário desenvolvedor de métodos de reconhecimento/casamento uma interface padrão para casamento e outra para reconhecimento. Assim ele pode realizar o desenvolvimento do método em seu ambiente local e após a finalização, submeter o método para avaliação dos administradores. Após essa avaliação, caso não haja problemas, o método será integrado ao sistema, possibilitando ao usuário utilizar toda a infraestrutura do sistema para testes do método ora desenvolvido.

- **Gerenciamento de usuários**

Nesta funcionalidade, o administrador do sistema pode gerenciar as contas dos usuários, as possíveis ações de gerenciamento são: ativação/inativação de usuário e alteração de perfil de acesso ao sistema.

3.6 Usuários do sistema

Dada a existência de uma série de funcionalidades dentro do sistema, é natural que certas funcionalidades estejam restritas a um certo conjunto de usuários. Por exemplo, a funcionalidade de validação do casamento/reconhecimento, deve ser realizada por um usuário com conhecimentos, que confirme se há erros no casamento/reconhecimento a fim de garantir a consistência dos dados de *ground-truth*.

Dada esta necessidade, o **ExpressMath Online** implementa um acesso controlado por perfis de usuário, cada um com acesso à um determinado conjunto de funcionalidades do sistema. A seguir iremos descrever quais são os perfis do sistema, assim como quais funcionalidades estão disponíveis no perfil e a motivação de sua criação.

1. **Voluntários**

É o perfil básico do sistema, no qual o usuário é capaz apenas de realizar a transcrição das expressões, executar o casamento/reconhecimento das mesmas (apenas para visualizar o resultado, não há funcionalidade de validação de EMs), além de ter a possibilidade de visualizar os resultados atuais do sistema.

2. Gerenciadores de EMs

É o perfil dedicado aos pesquisadores da área de reconhecimento de EMs. Neste perfil, o usuário é capaz de criar conjunto de expressões modelo, editar as parametrizações dos métodos caso existam, submeter uma nova requisição de teste do sistema, visualizar e realizar análise dos resultados, além de validar e corrigir o casamento/reconhecimento das expressões.

3. Desenvolvedores de métodos de rotulação

É o perfil que possibilita a adição de novos métodos de casamento ao sistema. Neste perfil, o usuário é capaz, através de uma interface de submissão, adicionar um novo método ao **ExpressMath Online**. O método ora submetido passa por uma avaliação dos administradores e é adicionado ao sistema caso seja aceito.

Além dessa funcionalidade específica, os desenvolvedores também tem acesso as funções de criação de conjunto de dados modelo, submissão de um novo teste de sistema, editar as parametrizações dos métodos caso existam, visualizar e realizar análise dos resultados, além de validar e corrigir o casamento/reconhecimento das expressões.

4. Administradores

É o super usuário do sistema. Ele possui todas as funcionalidades antes mencionadas, além de possuir acesso integral ao controle de usuários, ou seja, é ele quem atribui um determinado perfil à um usuário.

Para facilitar o entendimento, segue na tabela 3.1 a matriz de acessos às funcionalidades do sistema por perfil.

Funcionalidade	Perfil			
	1	2	3	4
Coleta de amostras de EMs baseadas em modelo	X	X	X	X
Coleta de EMs	X	X	X	X
Parametrizações		X	X	X
Processo de casamento de amostras de EMs baseadas em modelos		X	X	X
Criação de modelos de EMs		X	X	X
Teste de métodos de reconhecimento/casamento (Offline)		X	X	X
Teste de métodos de reconhecimento/casamento (Online)		X	X	X
Avaliação de expressões		X	X	X
Correção de expressões		X	X	X
Processo de Adição de novos métodos			X	X
Gerenciamento de usuários				X

Tabela 3.1: Matriz de acesso às funcionalidades por perfil. Na tabela, os perfis são descritos por números, onde respectivamente significam: Voluntários, Gerenciadores de EMs, Desenvolvedores de métodos de rotulação e Administradores.

Capítulo 4

Método para casamento de expressões

Como descrito no capítulo 2, o problema de casamento de expressões considera o mapeamento de símbolos de uma expressão transcrita aos correspondentes símbolos na expressão modelo. Uma vez realizado o casamento, informações de gabarito da expressão modelo podem então ser automaticamente transferidos para a expressão transcrita. O método de casamento de expressões proposto por Honda (Honda, 2013) impõe uma restrição severa ao supor que os símbolos encontram-se previamente segmentados em ambas as expressões, modelo e transcrição. Por outro lado, permite a modelagem do problema como um problema simples de casamento símbolo-a-símbolo.

Porém, muitas vezes essa restrição não pode ser suposta. Por exemplo, isso acontece nos casos de dados capturados sem a indução da segmentação durante a captura como feito no método proposto em Honda (Honda, 2013) ou pelo simples fato dessa suposição ser considerada muito restritiva.

Dadas essas características e implicações, neste trabalho propomos um outro método que realiza o casamento entre as expressões, porém sem considerar que os símbolos da transcrição estejam segmentados. Sem essa suposição, não podemos mais estabelecer uma relação direta entre as expressões, uma vez que a quantidade de elementos que compõem as expressões não são mais necessariamente iguais, já que agora os elementos básicos que compõem a transcrição são os traços e não os símbolos.

O problema de associação que estamos considerando pode ser visto como um problema de segmentação de símbolos, que por sua vez é um problema de particionamento do conjunto de traços que formam a expressão, seguida da identificação dos símbolos. Uma busca exaustiva sobre o conjunto de todas as possíveis partições de traços da expressão transcrita exigiria um tempo de processamento impraticável. Portanto, é necessário reduzir a quantidade de partições a serem examinadas.

Abordagens propostas na literatura para a segmentação de símbolos consideram algumas restrições para reduzir o número de partições a serem examinadas Koschinski *et al.* (1995); Matsakis (1999); Winkler e Lang (1997). Em geral, tais restrições são implemen-

tadas restringindo-se os agrupamentos de traços a serem examinados. A determinação dos agrupamentos a serem examinados pode explorar a informação temporal associada aos traços, assim como informações de características espaciais e visuais dos elementos presentes na expressão.

Em nosso problema podemos, adicionalmente, explorar o fato da expressão a ser segmentada ser uma transcrição de uma expressão modelo na qual os símbolos presentes são conhecidos. Ou seja, a expressão modelo pode ser utilizado como um guia para a segmentação e associação dos símbolos segmentados aos correspondentes símbolos no modelo.

Antes de detalhar o método desenvolvido, descrevemos inicialmente os pré-processamentos aplicados nas normalizações.

4.1 Pré-processamentos de EMs

De forma similar ao proposto no trabalho de Honda, são aplicados três pré-processamentos visando um registro espacial entre as duas expressões a serem casadas.

4.1.1 Representação de expressões por grafos

Para capturar a estrutura, expressões são representadas por meio de grafos com atributos. No caso das expressões modelo, os símbolos são representados por vértices, e a eles associa-se coordenadas do centro da caixa envoltória do correspondente símbolo. Arestas são incluídas ligando cada vértice aos demais vértices no grafo (resultando, portanto, em grafos completos). Atributos como comprimento e direção das arestas que saem de um vértice fornecem informação sobre a relação espacial do respectivo símbolo aos demais símbolos na expressão. A cada vértice fica associado o símbolo correspondente e portanto qualquer característica do mesmo pode ser obtida.

Em relação ao trabalho de Honda, a representação utilizada aqui difere em termos de vértices. Os vértices no caso de expressões transcritas correspondem aos traços e não aos símbolos.

4.1.2 Normalização espacial e de escala

O método de casamento no trabalho de Honda executa um registro prévio das duas expressões a serem casadas. Os grafos correspondentes à expressão modelo e à expressão transcrita (entrada) podem diferir em escala e em posição, dependendo do tamanho dos símbolos e da posição na qual as mesmas foram escritas. Desta forma, realiza-se a normalização de escala e posição da seguinte forma:

1. Obtém-se o centróide de ambos os grafos a partir de seus *bounding boxes*.

2. O grafo modelo é transladado de forma que as coordenadas de seu centróide coincidam com as coordenadas do centróide do grafo de entrada.
3. A largura do *bounding box* do grafo modelo é alterada de forma a coincidir com a largura do *bounding box* do grafo de entrada. Neste processo, os vértices do grafo modelo são também transladados proporcionalmente.

Note que no método de Honda não é importante qual dos grafos, modelo ou de entrada, é transladado e tem a escala ajustada ao outro. No entanto, no caso do problema tratado neste trabalho, uma vez que os elementos que compõem a expressão transcrita são os traços, a translação dos traços poderia “desconfigurar” símbolos. Assim, convém alterar o grafo modelo para que ele se ajuste à posição e escala do grafo de entrada (e não o contrário).

4.1.3 Normalização dos pontos do contorno do símbolo

Outra normalização realizada é a normalização da distribuição de pontos em um traço. Trata-se da mesma normalização descrita na seção 2.2. Os principais passos desse processo de normalização são transcritos aqui por conveniência:

1. **Passo 1:** Remoção dos pontos duplicados.
2. **Passo 2:** São filtrados os ruídos com uma técnica de *smoothing*.
3. **Passo 3:** A distância entre os pontos são equalizadas.

Na figura 4.1 é apresentado um exemplo de entrada e saída dessa técnica.

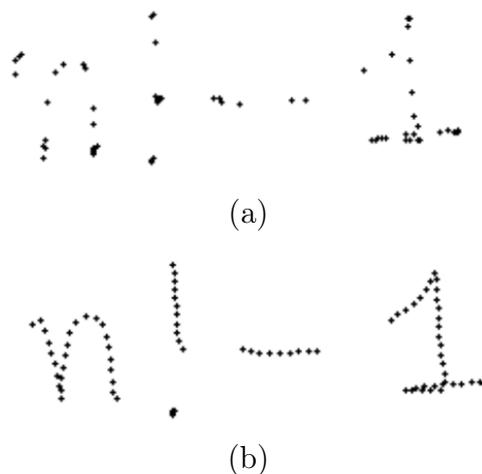


Figura 4.1: A figura apresenta a entrada e saída (a e b respectivamente) para o processo de normalização dos pontos do contorno do símbolo para a expressão $n! - 1$.

4.2 Casamento de expressões sem a suposição de segmentação prévia

Como mencionamos anteriormente, uma das contribuições deste trabalho é o desenvolvimento de um método para o casamento de expressões que não parte do pressuposto que os símbolos estejam corretamente segmentados. Inicialmente, diversas abordagens para o casamento de EMs baseadas em ideias utilizadas no problema de segmentação de símbolos (Koschinski *et al.*, 1995; Matsakis, 1999) foram pensadas, porém todas apresentavam um alto grau de complexidade e se mostravam ineficazes quando submetidas a testes na prática. Com isso, ficou evidente que as técnicas deveriam ser alteradas e tratadas de uma maneira mais simples e direta.

Um ponto importante a ressaltar antes de começarmos a descrição da técnica desenvolvida é a maneira na qual a expressão modelo é utilizada para realizar o casamento. As abordagens iniciais partiam da transcrição e tentavam encontrar os correspondentes na expressão modelo. A alteração principal da nova técnica é a utilização da expressão modelo como guia para a realização do casamento entre as expressões.

Assim, o novo método tenta, para cada símbolo da expressão modelo, encontrar os traços correspondentes na transcrição. Esse processo está subdividido em cinco etapas principais, sendo elas: seleção de traços candidatos, cálculo de verosimilhança, determinação de uma associação preliminar, remoção de associações sobrepostas e processamento de símbolos e traços órfãos. A descrição das cinco etapas, além da forma como elas interagem entre si e com a expressão modelo são apresentadas a seguir.

Desse método é esperado a produção de uma partição dos traços da transcrição, com cada símbolo da expressão modelo associado a uma das partes da partição.

4.2.1 Etapa 1: Seleção de traços candidatos

Para cada símbolo s_i da expressão modelo, desejamos determinar o conjunto de traços da expressão transcrita que formam o símbolo correspondente a s_i . Note que devido ao pré-processamento de normalização espacial e de escala descrito na seção 4.1.2, ambas as expressões estão “registradas” e portanto é esperado que as posições ocupadas pelos traços desejados na expressão transcrita sejam relativamente próximas ao do símbolo s_i .

Desta forma, para cada símbolo s_i presente na expressão modelo, inicialmente associamos um conjunto $T(s_i)$ de até 6 traços da transcrição. A escolha desse número, 6, foi feita de forma empírica, levando-se em conta o fato de a maior parte dos símbolos manuscritos possuir, no máximo, 4 traços, e também para manter controle sobre a complexidade computacional. Além disso, traços que encontram-se relativamente distantes uns dos outros não devem provavelmente fazer parte de um único símbolo. Portanto, para construir esse conjunto $T(s_i)$,

toma-se os traços da transcrição cujo centróide esteja a uma distância máxima d do centróide de s_i . Essa distância d é definida como o comprimento da diagonal da caixa envoltória de s_i multiplicado por um fator definido pela equação 4.1:

$$fd = \frac{w_t}{w_m} \quad (4.1)$$

onde w_t e w_m são respectivamente os tamanhos da expressão transcrita e modelo. Na figura 4.2 é apresentada uma ilustração do processo de seleção de traços.

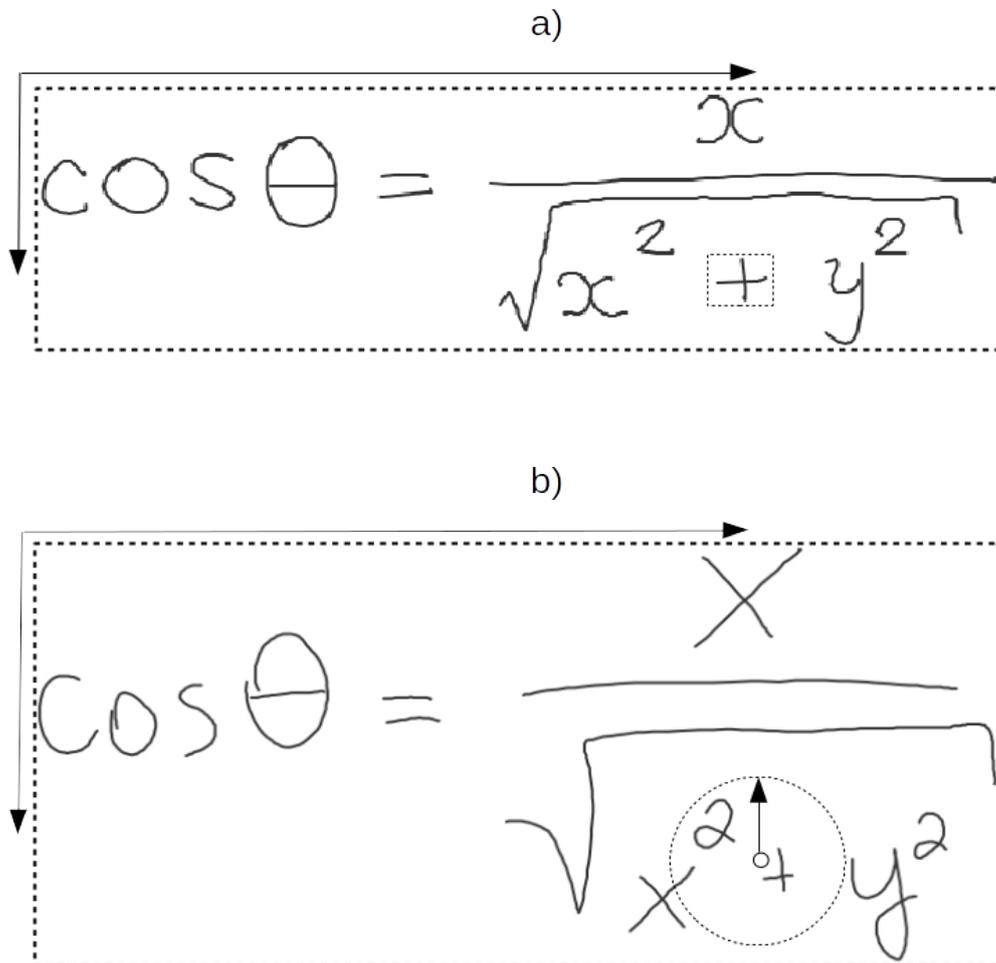


Figura 4.2: Ilustração do processo de seleção de traços. Em a, apresentamos a posição relativa ao símbolo x da expressão $\cos \theta = \frac{x}{\sqrt{x^2 + y^2}}$; em b apresentamos o centróide de x sobreposto na expressão transcrita, juntamente com disco definido pela distância considerada para a construção do conjunto de traços que podem corresponder a x .

Com isso, temos três hipóteses para $T(s_i)$: um conjunto dentro do limite de seis traços, um conjunto sem nenhum traço e um conjunto com mais de seis traços. A primeira hipótese não precisa de pós tratamento, uma vez que já está dentro das condições estabelecidas. Na segunda hipótese, como não foram selecionados traços, a distância de corte mencionada no processo anterior é aumentada em 10% e o processo de seleção é repetido. Caso novamente

não seja selecionado nenhum traço, a distância é aumentada novamente e assim sucessivamente até que o conjunto de traços selecionado seja não vazio. Por fim, na terceira hipótese, em que são selecionados mais que seis traços da transcrição, faz-se necessário algum critério de seleção adicional dentro desse conjunto. Para isso, selecionamos os seis traços mais próximos ao centróide de s_i .

Vale observar que na segunda hipótese, no decorrer das seguidas repetições do processo de seleção, pode ocorrer de serem selecionados mais que seis traços para o conjunto. Neste caso, o conjunto passa a se enquadrar na terceira hipótese e passa pelo processamento mencionado acima.

4.2.2 Etapa 2: Cálculo de verosimilhança

Após a etapa de seleção de traços, temos um conjunto de componentes (traços) candidatos $T(s_i)$ para cada símbolo s_i presente na expressão modelo. Para cada símbolo s_i da expressão modelo, testamos possíveis combinações dos traços em $T(s_i)$ para determinar qual delas melhor se assemelha ao símbolo s_i . Para isso, um classificador de símbolos baseado em redes neurais desenvolvido em [Julca-Aguilar *et al.* \(2014\)](#) e que apresenta ótimos resultados é aplicado sobre cada combinação de traços.

O classificador acima mencionado possui uma característica interessante, pois além de reconhecer se um conjunto de traços corresponde a um símbolo, ele também reconhece agrupamentos que não correspondem a nenhum símbolo. Essa característica é interessante pois pode ser utilizada para rejeitar algumas combinações de traços. O classificador é treinado para reconhecer classes de símbolos de um conjunto pré-estabelecido, incluindo também uma classe de rejeição (não-símbolos). Ele recebe como entrada um conjunto de traços e, para cada classe de símbolos, devolve um grau de pertinência daquele conjunto de traços à classe. Frequentemente há uma classe dominante, indicando que muito provavelmente o conjunto de traços corresponde àquela classe. Dessas, o que nos interessa é o grau de pertinência à classe do símbolo s_i .

Custo de associação entre símbolos

Para computar o custo de associação entre um símbolo s_i da expressão modelo e uma combinação de traços $T_j \subseteq T(s_i)$ da expressão transcrita, além de utilizarmos o classificador de símbolos, utilizamos informação de distância.

O resultado do classificador de símbolos é usado para definir o custo de associação de símbolo, de acordo com a equação 4.2, na qual $c_c(s_i, T_j)$ é o grau de confiança devolvida pelo classificador de símbolos quanto a T_j ser da mesma classe de s_i .

$$c_a(s_i, T_j) = 1 - c_c(s_i, T_j). \quad (4.2)$$

Assim, se o grau de confiança é alto (próximo de 1, indicando que muito provavelmente os traços correspondem ao símbolo s_i), o correspondente custo é baixo (próximo de zero).

Além disso, um custo associado à distância é definido como a razão entre a distância d_t entre o centróide do símbolo s_i e o do conjunto de traços T_j considerado e um fator de normalização D_p , denominada diagonal ponderada da transcrição. Este custo denominado aqui por custo de distância é definido pela equação 4.3:

$$c_d = \frac{d_t}{D_p} \quad (4.3)$$

A diagonal ponderada D_p é definida por

$$D_p = \frac{D_t}{|S_m|} \quad (4.4)$$

na qual D_t é a diagonal do retângulo envolvente da transcrição e $|S_m|$ é a quantidade de símbolos que compõem a expressão modelo. Essa diagonal tem o objetivo de definir uma distância "padrão" para a expressão transcrita.

Finalmente, o custo de associação entre s_i e T_j é definido pela seguinte equação:

$$C(s_i, T_j) = \alpha c_a(s_i, T_j) + (1 - \alpha)c_d. \quad (4.5)$$

O parâmetro $\alpha \in [0, 1]$ é um fator de ponderação entre o custo de associação de símbolo e o custo de distância. O valor de α desta equação foi definido empiricamente e fixado em 0.9 nos testes realizados.

Em termos qualitativos, este custo leva em consideração o grau de confiança do classificador e a distância entre os centróides do símbolo s_i e do conjunto de traços T_j selecionado para ser associado a ele.

4.2.3 Etapa 3: Associação preliminar

Após computados os custos de associação entre s_i e as combinações do conjunto $T(s_i)$, escolhe-se inicialmente a combinação $T_j \subseteq T(s_i)$ com melhor custo de associação.

Uma vez que para cada símbolo s_i o conjunto de traços candidatos $T(s_i)$ e uma combinação $T_j \subseteq T(s_i)$ de melhor custo é escolhido independentemente, um mesmo traço pode ser associado a mais de um símbolo da expressão modelo.

Outro ponto importante que devemos mencionar é que nos casos em que o reconhecedor de símbolos não aceita nenhuma das combinações como sendo da mesma classe de s_i , o símbolo é marcado como não associado e ele é deixado de lado para que seja tratado posteriormente.

4.2.4 Etapa 4: Remoção de associações sobrepostas

Na etapa de determinação de uma associação preliminar escolhemos, para cada símbolo, a combinação de traços que resulta em melhor custo de associação dentre as combinações de traços analisadas. Como na construção do conjunto de traços candidatos, a partir do qual as combinações são geradas, não nos preocupamos se um traço está presente em mais de um conjunto de traços candidatos, podem ocorrer casos em que um mesmo traço esteja em duas combinações escolhidas no processo de associação preliminar. Isto configura um erro de casamento e deve ser reexaminado.

Para isso, dentre as associações que envolvem um traço t em comum, selecionamos aquela com melhor custo de associação. As demais associações que envolvem o traço t são descartadas para serem tratadas posteriormente.

4.2.5 Etapa 5: Processamento de símbolos e traços órfãos

Como descrito acima, podem ocorrer casos onde não são encontradas associações para determinados símbolos, resultando no que chamamos de símbolos órfãos. Analogamente, pode ocorrer também casos onde traços não foram associados a nenhum símbolo na expressão modelo, resultando no que chamamos de traços órfãos.

Ambos casos mencionados necessitam de pós-processamentos para serem corretamente associados. Para isso, realizamos dois pós-processamentos, um para associar os traços órfãos aos símbolos órfãos e outro que tenta associar traços extras às soluções encontradas nas etapas anteriores. A seguir, descreveremos como é realizado o processamento para cada um dos casos.

Remoção de símbolos órfãos

Para fazer associações aos símbolos órfãos, é necessário que também existam traços órfãos. Para cada símbolo órfão s_i , é criado um conjunto com até seis traços órfãos mais próximos ao centróide de s_i . O processo é o mesmo realizado nas etapas anteriores, porém agora apenas os traços órfãos são considerados como candidatos e não se impõe uma restrição de distância. Assim, combinações que não haviam sido testadas podem ser selecionadas aumentando a resiliência do método.

Remoção de traços órfãos

O fato de existirem traços órfãos, evidencia que o conjunto de traços associados a um ou mais símbolos da expressão modelo estão incorretos. Portanto é necessário que esses traços passem a fazer parte do grupo de traços associado a algum símbolo.

Para incorporar os traços órfãos a uma associação, criamos para cada traço órfão uma

lista com as associações que envolvem os símbolos mais próximos ao centróide do traço (para computar essa distância, utilizamos o centróide do *bounding box* do conjunto de traços associados aos símbolos). Em seguida, testamos as associações ora encontradas, agora adicionando o traço órfão. Caso alguma das novas associações tenha custo menor do que a associação anterior, escolhemos a de menor custo e incorporamos o traço órfão ao conjunto de traços da associação escolhida. Porém, se acrescentar o traço a qualquer das associações existentes não melhorar o custo, o traço órfão fica sem associação.

Esse processo de remoção de símbolos e traços órfãos não garante que todos eles serão eliminados. Do processo podem restar símbolos sem associação ou traços órfãos e esses configurarão erro no casamento. Uma associação entre um símbolo s_i na expressão modelo e um conjunto de traços T_j na expressão transcrita está correta se o conjunto T_j contém exatamente os traços do símbolo correspondente a s_i (i.e., T_j não deve conter traços de outro símbolo e todos os traços do símbolo correspondente a s_i na expressão transcrita devem estar em T_j).

4.3 Resultados experimentais

O método desenvolvido foi testado sobre o conjunto de dados do ExpressMatch (Julca-Aguilar e Hirata., 2012), e apresentou taxa de acerto de associação de símbolos maiores que 79%. Os resultados, separados por expressões modelo, são apresentados na figura 4.3.

Esse resultado foi obtido utilizando a funcionalidade de avaliação de resultados que compõe o ExpressMath Online. Na seção 5.5 são apresentados maiores detalhes dessa avaliação.

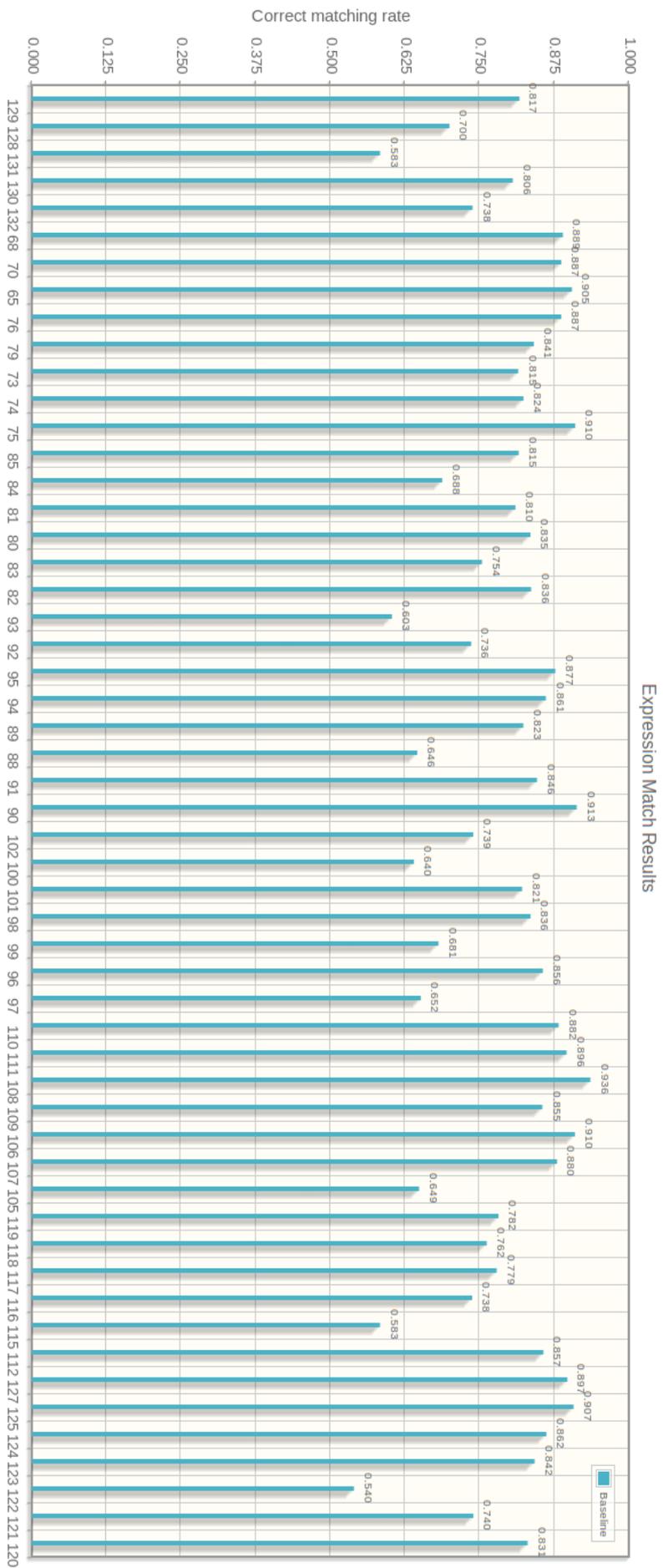


Figura 4.3: No gráfico, são apresentados as taxas de acerto de associação de símbolos separados por expressão modelo.

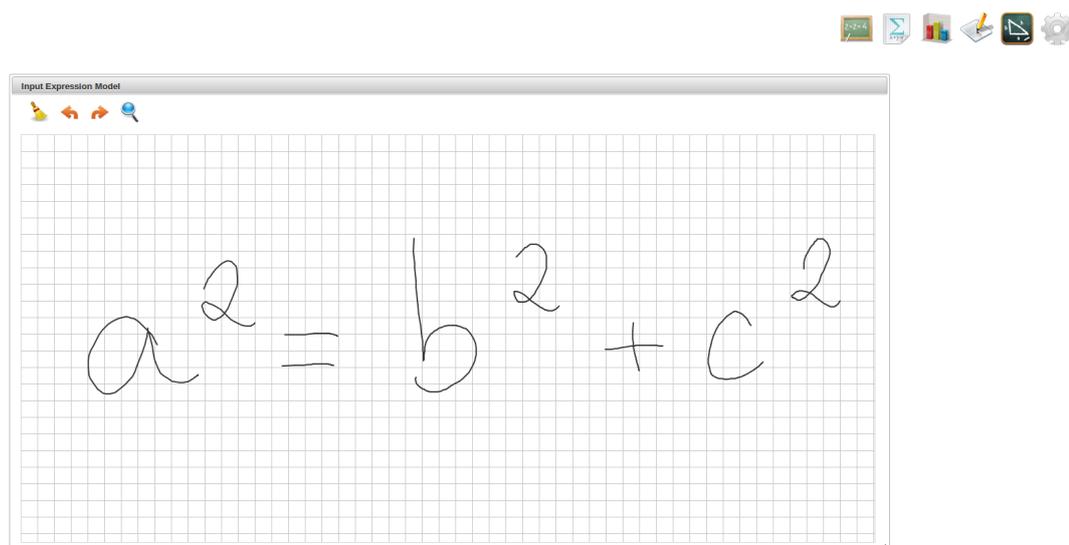
Capítulo 5

ExpressMath Online na prática

Neste capítulo, apresentaremos o sistema **ExpressMath Online** na prática. Iremos demonstrar como o sistema desempenha um determinado conjunto de funcionalidades e expor as idéias que basearam a construção do sistema dessa maneira. Além disso, na sessão 5.5, aprofundaremos a análise dos resultados obtidos com o método de casamento apresentado no capítulo anterior.

5.1 Criação dos modelos

Para a criação dos modelos dentro do ExpressMath Online, inicialmente o usuário insere a expressão modelo que deseja. Em seguida, rotula manualmente os símbolos da expressão atribuindo a cada um deles o respectivo gabarito, além de informar o gabarito da expressão modelo em formato $\text{L}^{\text{T}}\text{E}^{\text{X}}$.



Welcome davi - [Logout](#)

VISION science 

Figura 5.1: Painel de inserção de expressões modelo.

Após a inclusão, a expressão modelo estará automaticamente disponível no painel de

transcrições e já será possível coletar transcrições deste novo modelo. Na figura 5.1 é apresentado o painel de coleta das expressões modelo.

Vale ressaltar que a funcionalidade de coleta de modelos encontra-se em desenvolvimento e ainda não é possível realizar a rotulação dos símbolos, apenas a rotulação da expressão.

5.2 Transcrição das expressões

Dentre as funcionalidades que compõem o **ExpressMath Online**, uma das mais citadas ao decorrer deste trabalho, é a transcrição das expressões modelo, que tem a finalidade de aumentar significativamente a volumetria dos dados de EMs online.

A funcionalidade é composta por dois painéis, como apresentado na figura 5.2. Na figura, o primeiro painel a esquerda, é o painel das expressões modelo, no qual aparecem as expressões que devem ser transcritas. No segundo painel, à direita, é onde as expressões devem ser transcritas.

Na figura 5.3, temos as oito opções principais do processo de transcrição de expressões. A seguir descreveremos o que cada uma delas desempenha dentro do processo.

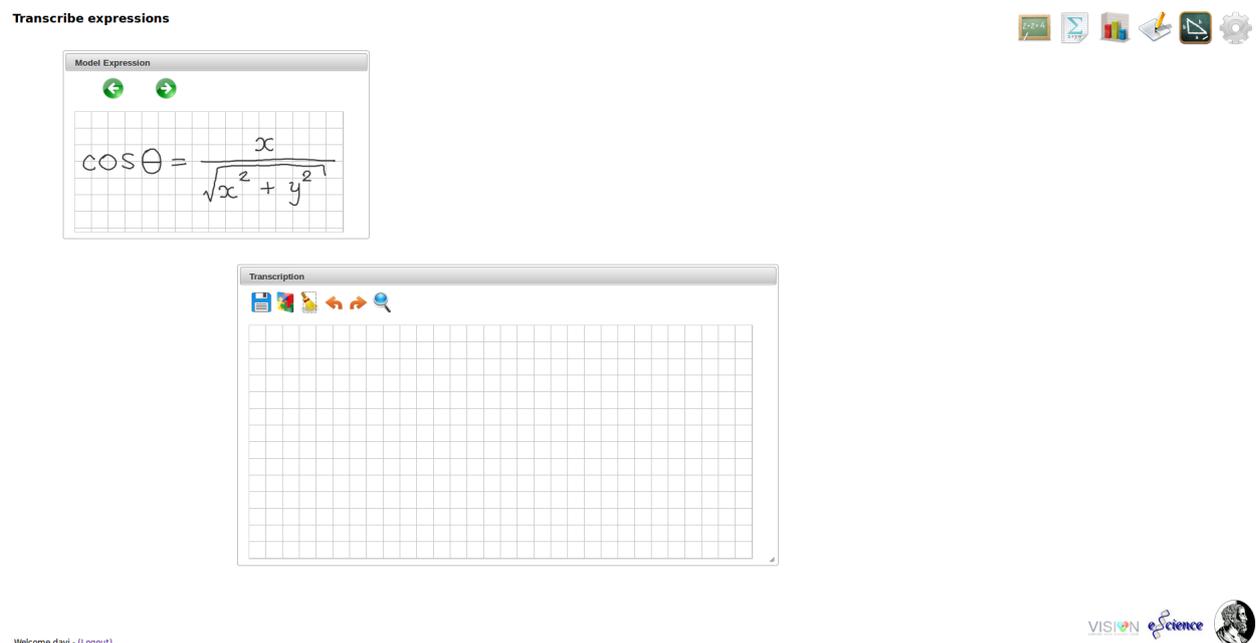
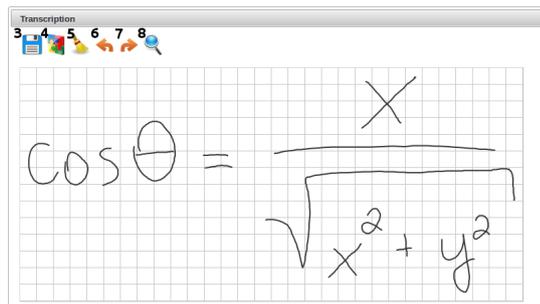
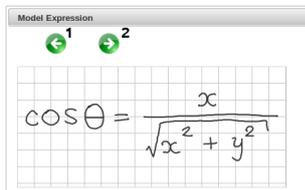


Figura 5.2: Funcionalidade de transcrição das expressões modelo. O primeiro painel, localizado na parte superior esquerda da figura é onde são apresentadas as expressões modelo, e no segundo painel é onde as transcrições são coletadas.

1. **Expressão modelo anterior:** navega para a expressão modelo anterior.
2. **Próxima expressão modelo:** navega para a próxima expressão modelo.
3. **Salvar:** salva a transcrição. Esse é o ponto inicial do fluxo de vida das EMs para o caso de casamento.

4. **Testar casamento:** realiza um casamento entre a transcrição e a expressão modelo. Nesse processo não há alterações dentro do fluxo de vida das expressões, ele existe para o caso do usuário querer ver o casamento da transcrição sem fazer acesso em outros menus.
5. **Limpar:** remove todos os traços da transcrição
6. **Desfazer traço:** remove o último traço incluído à transcrição.
7. **Refazer traço:** refaz o último traço excluído da transcrição.
8. **Zoon:** aumenta/diminui o tamanho da transcrição.

Transcribe expressions



Welcome davi - [Logout](#)



Figura 5.3: Funcionalidade de transcrição das expressões modelo, após a transcrição ter sido coletada.

5.3 Teste de métodos de casamento

Como descrito na seção 3.5, uma das funcionalidades presentes no **ExpressMath Online** é o teste de métodos de casamento das transcrições com seus respectivos modelos. Essa funcionalidade tem grande importância dentro o grupo de funcionalidades do ExpressMath Online, pois ela agiliza o trabalho dos pesquisadores na análise de problemas pontuais no casamento. Ele oferece, por exemplo, uma maneira semi-automática de realizar a análise de problemas dos métodos em determinadas amostras dos dados. Na figura 5.4 é apresentado o painel onde o teste dos métodos de casamento é realizado.

Por exemplo, em um teste hipotético que envolve grande volume de dados e problemas pontuais em determinadas transcrições, para realizar a análise dos problemas, geralmente os

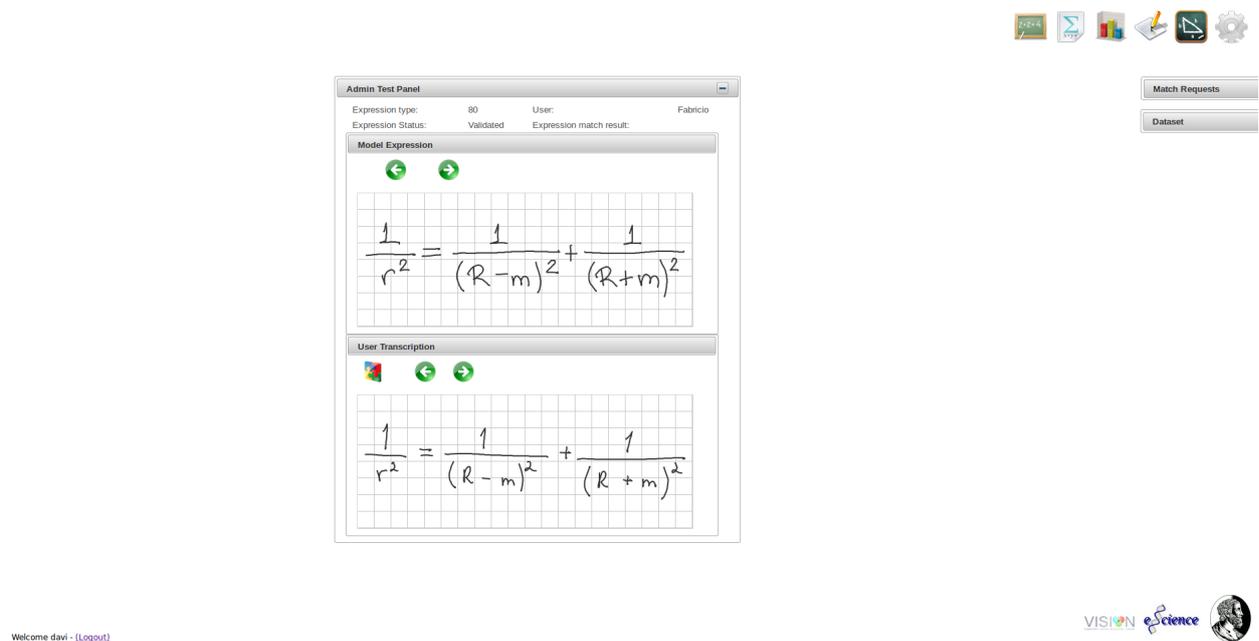


Figura 5.4: Painel de teste de casamento entre expressões modelo e transcrições. Nesse painel é possível observar o estado da expressão, além do usuário que transcreveu a expressão.

pesquisadores isolam os dados e aplicam os métodos e avaliam os resultados pontualmente de forma manual. Porém esse processo é lento e cansativo e pode demandar muito esforço dependendo da quantidade de dados a serem analisados.

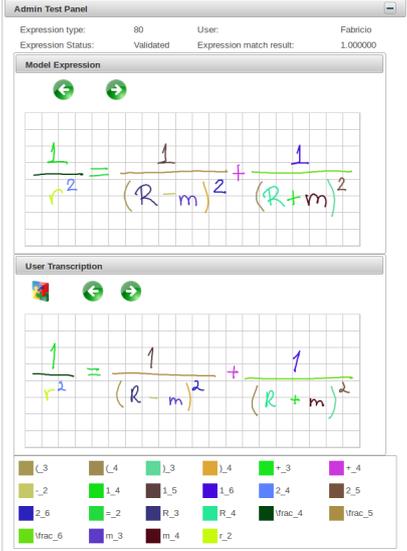
Dentro da funcionalidade de testes dos métodos de casamento do ExpressMath Online, o usuário poderá escolher qual método testar, e com quais dados quer realizar o teste, além de visualizar em tempo real os resultados dos casamentos, expressão por expressão.

Na figura 5.5 é apresentado o mesmo painel após o método de casamento ter sido aplicado. Vale notar que na segunda figura, os símbolos em ambas as expressões foram coloridos e através das cores é possível conferir o casamento e as classes dos símbolos casados. Além dessas informações, também é apresentado o percentual de acerto no casamento das expressões, assim como o usuário que a transcreveu.

5.4 Parametrizações

Uma das preocupações na construção do ExpressMath Online, é prover uma maneira simples de desempenhar testes em diferentes conjuntos de dados e com configurações diferentes. De forma geral, para realizar testes com diferentes configurações, o processo é realizado de uma maneira manual. Isso significa que se algum usuário desenvolvedor de métodos de casamento/reconhecimento queira mudar o valor de algum parâmetro da configuração, ele precisa necessariamente construir novamente a aplicação para que as novas configurações sejam consideradas.

Além disso, o processo de escolha da configuração dos métodos é geralmente longo, cansa-



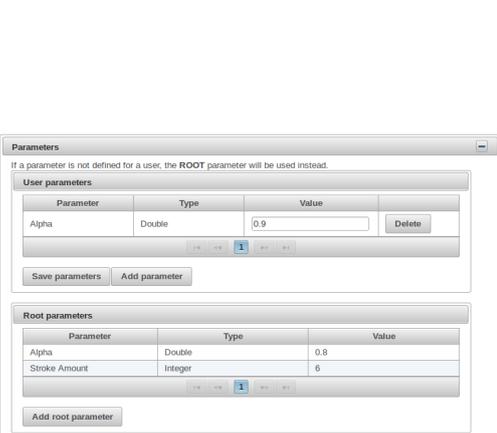
The screenshot shows the 'Admin Test Panel' interface. At the top, it displays 'Expression type: 80', 'User: Fabricio', 'Expression Status: Validated', and 'Expression match result: 1.000000'. Below this, there are two sections: 'Model Expression' and 'User Transcription'. Both sections show a handwritten mathematical expression: $\frac{1}{r^2} = \frac{1}{(R-m)^2} + \frac{1}{(R+m)^2}$. The 'Model Expression' section shows the expression with green and red highlights, and the 'User Transcription' section shows the same expression with blue and red highlights. Below the transcription, there is a color-coded legend for the parameters used in the expression, such as 'L_3', 'L_4', 'r_3', 'r_4', 'frac_3', 'frac_4', 'frac_5', 'frac_6', 'm_3', 'm_4', and 'r_2'.

Welcome davi - [Logout](#)



Figura 5.5: Painel de teste de casamento de expressões após o casamento ter sido realizado. Neste ponto é possível visualizar como o casamento ocorreu, além da taxa de acerto do casamento.

tivo e não necessariamente é realizado pelos desenvolvedores dos métodos. Contudo a abordagem descrita anteriormente impede que usuários não desenvolvedores testem as diferentes configurações dos métodos de casamento/reconhecimento, uma vez que o conhecimento em programação se torna imprescindível para a alteração das configurações.



The screenshot shows the 'Parameters' panel. It contains two sections: 'User parameters' and 'Root parameters'. The 'User parameters' section has a table with columns 'Parameter', 'Type', 'Value', and 'Delete'. The table contains one row: 'Alpha', 'Double', '0.9', and a 'Delete' button. Below the table are 'Save parameters' and 'Add parameter' buttons. The 'Root parameters' section has a similar table with columns 'Parameter', 'Type', and 'Value'. The table contains two rows: 'Alpha', 'Double', '0.8' and 'Stroke Amount', 'Integer', '6'. Below the table is an 'Add root parameter' button.

Welcome davi - [Logout](#)



Figura 5.6: Painel de parametrizações do sistema. Nele é possível definir os valores dos parâmetros utilizados nos métodos de casamento/reconhecimento.

Com isso, é necessário que uma abordagem diferente seja utilizada para fornecer um mecanismo mais dinâmico para que não só usuários com conhecimento em programação possam

testar os métodos. Portanto, dentro do ExpressMath Online temos uma funcionalidade específica para parametrizações.

A figura 5.6 apresenta o painel onde essas configurações são realizadas dentro do sistema. Nela, existem dois painéis, o primeiro, localizado na parte superior apresenta os parâmetros específicos do usuário. Já no segundo, são apresentados os parâmetros gerais do sistema, que são utilizados caso o usuário não tenha sobrescrito algum parâmetro.

Com isso, o sistema oferece a possibilidade para qualquer usuário personalizar as configurações dos métodos para realizar qualquer teste que desejar.

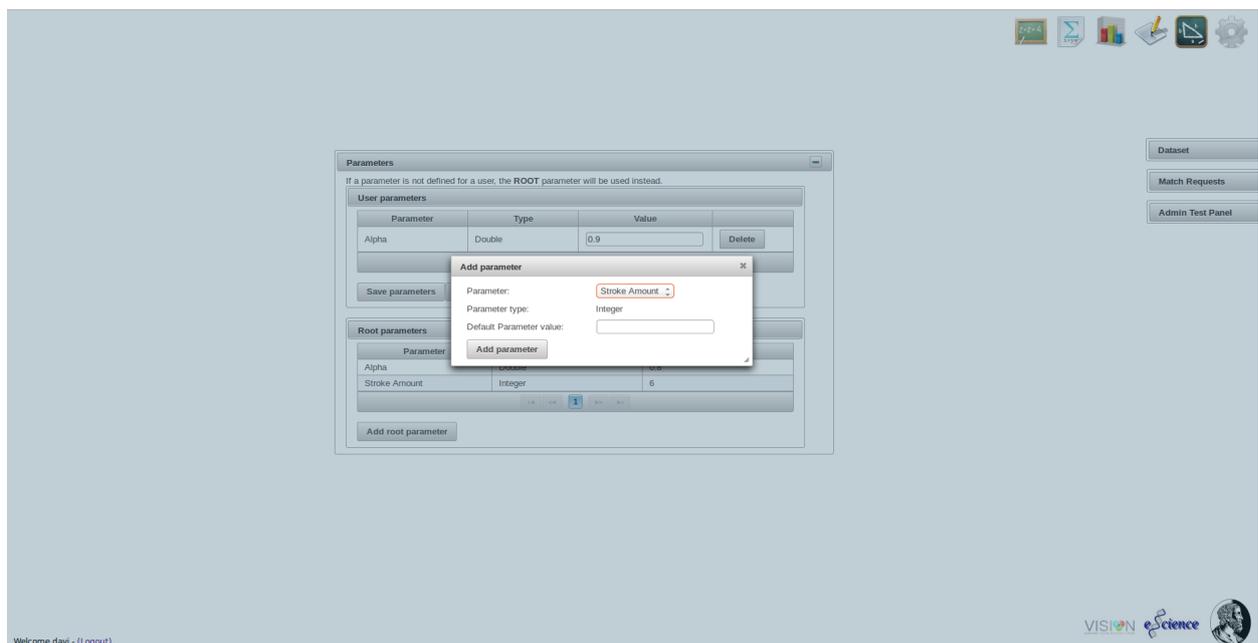


Figura 5.7: Opção de adição de um novo parâmetro de usuário dentro do ExpressMath Online. Nela o usuário escolhe o parâmetro que deseja definir além de seu valor.

Nas figuras 5.7 e 5.8 são apresentadas as funcionalidade de inclusão de parâmetros de usuário e gerais, respectivamente. No primeiro painel é apresentado ao usuário os parâmetros que podem ser configurados e seus respectivos tipos, e é requisitado o valor do parâmetro escolhido. No segundo painel é requisitado ao usuário o nome, descrição, tipo e valor padrão do novo parâmetro.

Vale ressaltar que imediatamente após a inclusão de um parâmetro geral dentro do ExpressMath Online, os usuários já podem realizar configurações próprias para os métodos. Essa nova configuração é realizada através da seleção do novo parâmetro adicionado, que irá aparecer no campo 'parameter' da figura 5.7.

Além das características apresentadas anteriormente, a parametrização dinâmica permite que usuários testem métodos não desenvolvidos por eles próprios com configurações próprias. Esta característica deverá contribuir diretamente para que o ExpressMath Online seja adotado pelos pesquisadores como ferramenta padrão para o tratamento de EMs.

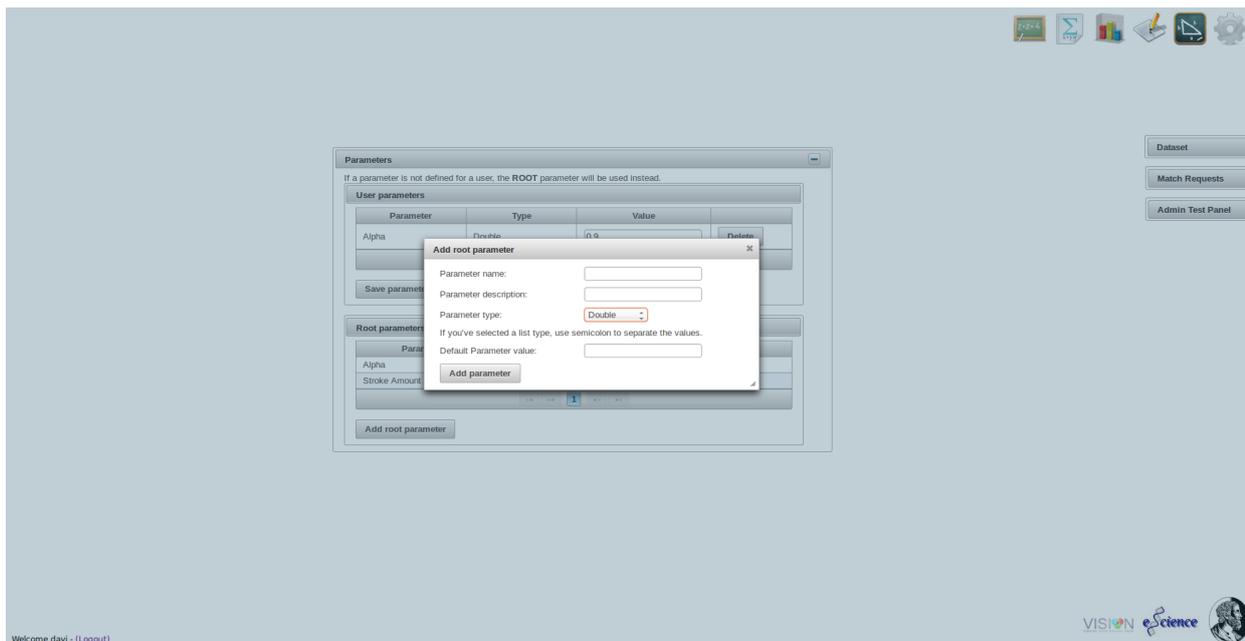


Figura 5.8: Opção de adição de um novo parâmetro padrão dentro do ExpressMath Online. Nela o usuário define o nome, descrição, tipo e valor padrão do novo parâmetro.

5.5 Avaliação dos resultados

Geralmente, em sistemas de tratamento de dados, existem mecanismos para a avaliação dos resultados de processamentos executados sobre os dados. Dada essa característica, dentro do ExpressMath Online existe uma funcionalidade para análise dos dados gerados pelo sistema.

Na figura 5.9, é apresentado o painel a partir do qual a configuração das análises são realizadas. Nele é possível selecionar quais dados estarão presentes no gráfico, além do tipo do gráfico desejado, de linhas ou barras.

A seguir, com a finalidade de expor o potencial da funcionalidade de avaliação dos resultados, apresentamos uma breve análise dos resultados do método apresentado no capítulo anterior. Primeiramente, como descrevemos acima, a funcionalidade de parametrizações é muito útil na análise dos dados, pois permite que diversas configurações sejam testadas dinamicamente. Dada essa característica, iremos realizar a análise do parâmetro α , presente no método de casamento. Essa análise é feita definindo diversos valores para a variável α , como apresentado nas figuras 5.10 e 5.11.

Após definidos os valores de α , é necessário fazer a análise do impacto da troca do valor do parâmetro no método de casamento. Essa análise é realizada através do painel de requisições que foi definido na sessão 3.5. Na figura 5.12, é apresentado o painel mencionado, onde é requisitado um nome e a motivação para a realização do teste. Note que no painel existem diferentes requisições para testar o método de casamento, uma parte delas é destinada a análise do parâmetro α , onde diversos valores foram definidos.

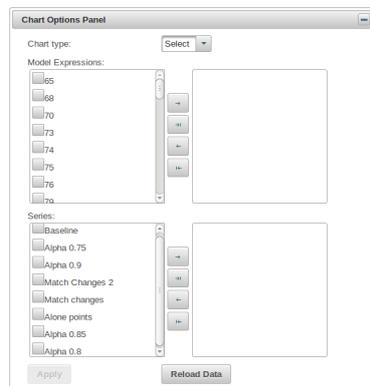
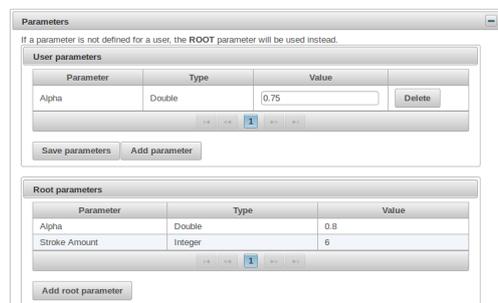


Chart Panel

Welcome davi - [Logout](#)



Figura 5.9: Painel de configuração das avaliações das requisições de casamento. Dentro deste painel é possível configurar o tipo do gráfico desejado, o conjunto de dados no qual será realizada a análise, além das séries disponíveis que irão compor o gráfico.



Match Requests
Admin Test Panel
Dataset

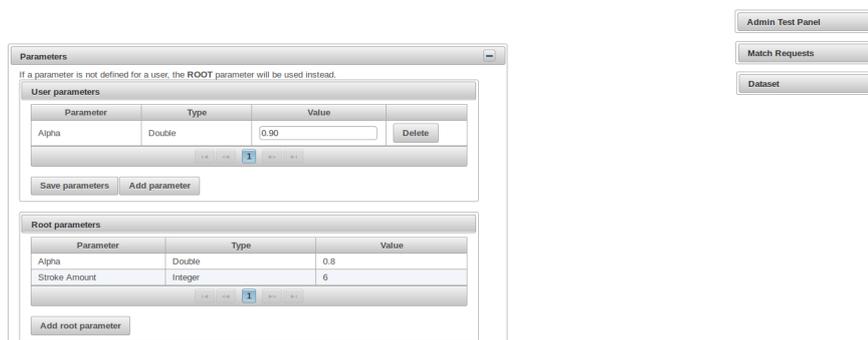
Welcome davi - [Logout](#)



Figura 5.10: Configuração do parâmetro α , definido como 0,75.

Após terem sido concluídas as requisições de teste, no painel de análise já estão disponíveis as séries para serem analisadas. Na figura 5.13 é apresentada a comparação entre duas das séries que haviam sido requisitadas. Note que na figura todos os dados foram utilizados, porém a análise poderia acontecer só em um determinado conjunto de dados, assim como a quantidade de séries presentes na análise.

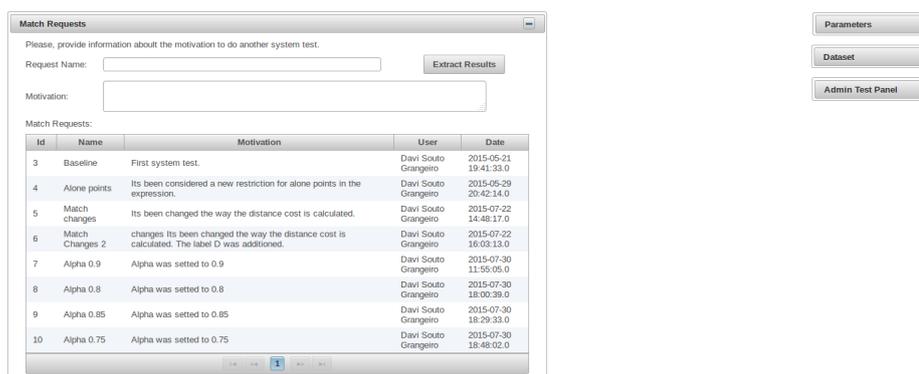
É importante destacar a praticidade do fluxo apresentado nesta sessão, pois ele permite



Welcome davi - [Logout](#)



Figura 5.11: Configuração do parâmetro α , definido como 0,9.



Welcome davi - [Logout](#)



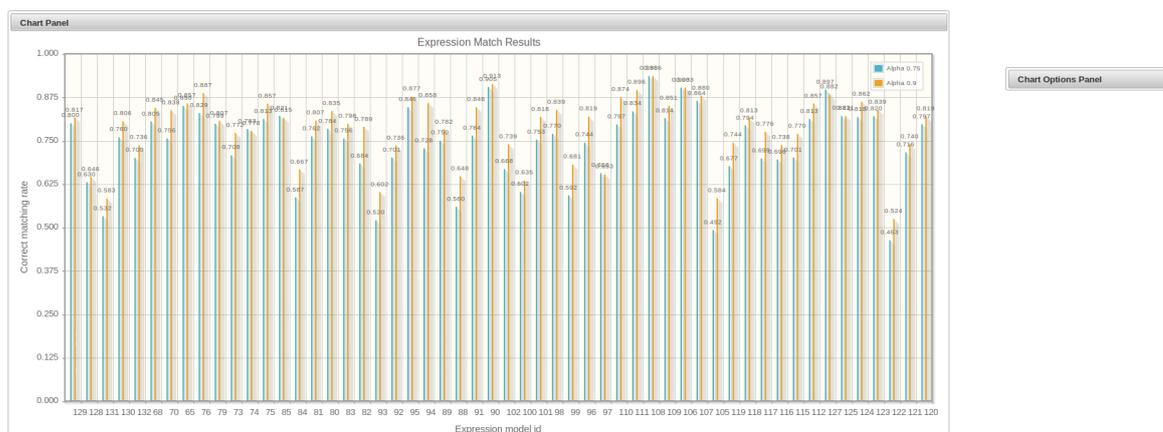
Figura 5.12: Painel de requisições de teste do ExpressMath Online. Vale notar que dentre as requisições inseridas, existe um conjunto delas que é destinado a análise do parâmetro α .

que análises complexas sejam realizadas de uma maneira trivial.

5.5.1 Erros de casamento mais comuns

Como vimos acima, algumas expressões não apresentaram resultado satisfatório no casamento. Com isso, uma análise se faz necessária a fim de evidenciar os pontos falhos da técnica apresentada.

Nota-se que alguns fatores influenciam diretamente o resultado do casamento, são eles:



Welcome davi - [Logout!](#)



Figura 5.13: Painel onde os gráficos de análise de resultados são apresentados. Esse gráfico é dinâmico, e suas informações são provenientes das configurações realizadas no painel apresentado na figura 5.9. No gráfico que há na figura, podemos observar a análise de duas requisições de casamento que tem a finalidade de observar o impacto da alteração do parâmetro α .

Associações Invertidas: Uma vez que o algoritmo de casamento proposto se baseia somente nas posições dos símbolos da expressão modelo para buscar os traços da transcrição, caso o registro espacial das duas expressões difira muito, podem ocorrer associações invertidas. Por exemplo, em uma expressão que possua dois símbolos próximos que pertençam a mesma classe. Na figura 5.14 apresentamos um casamento onde o erro mencionado ocorreu.

Espaçamento entre símbolos: algumas expressões apresentam um espaçamento desigual, influenciando diretamente na correspondência espacial entre as expressões modelo e transcrita. Uma expressão que possui essa característica é apresentada na figura 5.15.

Variação intrínseca da escrita: No caso de classificação baseada em protótipos, a variação intrínseca da escrita dificulta o processo de reconhecimento de símbolos. O fato de um mesmo símbolo poder ser escrito de diversas formas, conforme mostrado por exemplo na figura 5.16, influencia diretamente no desempenho do casamento.

Uma vez que uma nova forma de se escrever um determinado símbolo aparece, o classificador mencionado no capítulo anterior precisa necessariamente ser treinado novamente. Ao decorrer dos testes, algumas expressões apresentaram erro de casamento pois seus símbolos foram escritos de maneira diferente nas duas expressões, um exemplo é apresentado na figura 5.17.

Uma funcionalidade que será incorporada ao **ExpressMath Online** futuramente é o retreinamento do classificador ao decorrer do crescimento do banco de dados de expressões validadas. Esse retreinamento tende a melhorar o desempenho do classificador uma vez que,

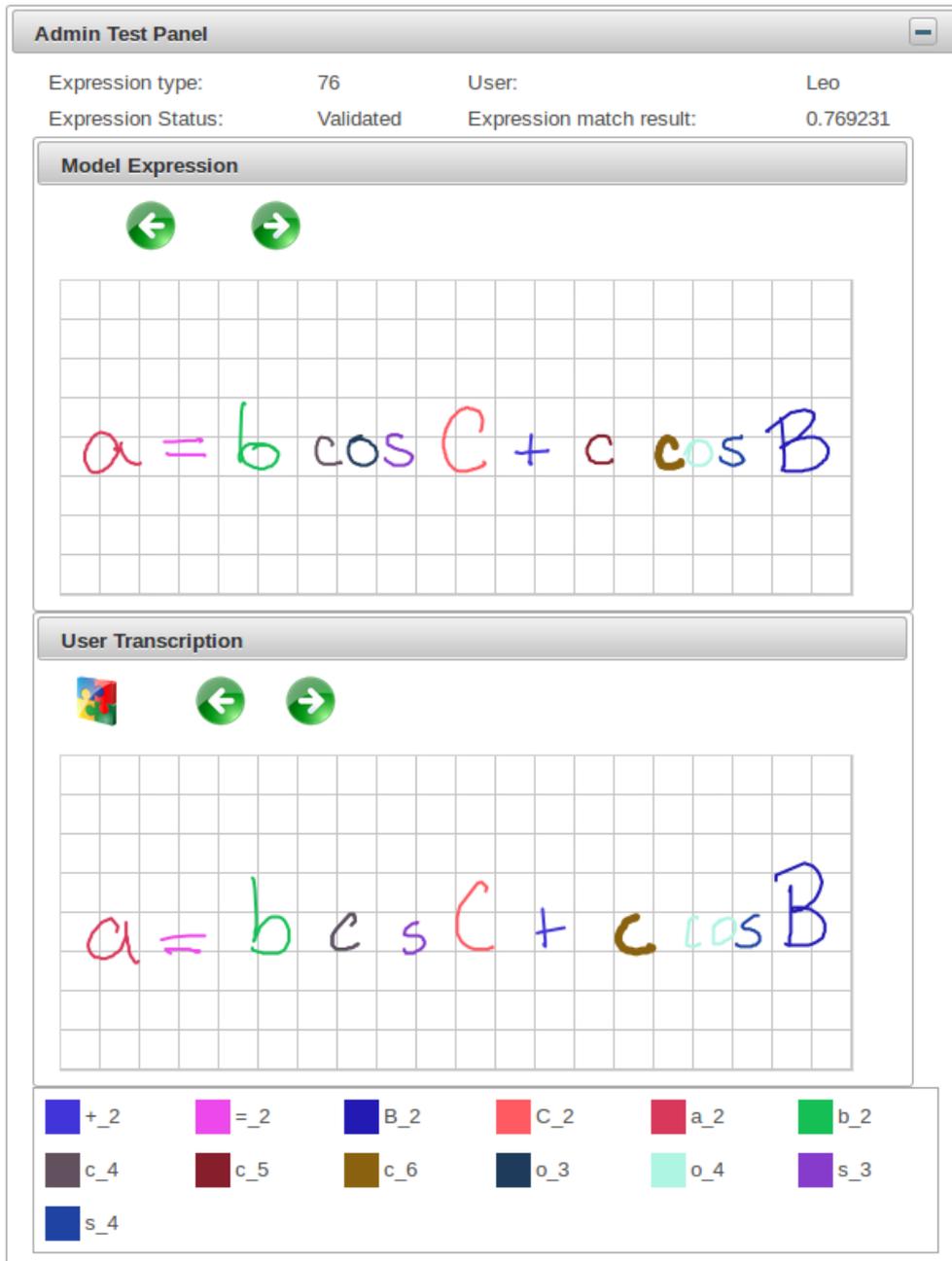


Figura 5.14: Caso de casamento de expressões onde ocorreu uma associação invertida entre duas letras 'c'.

com o decorrer do tempo, a base de dados aumenta e se torna mais heterogênea.

Símbolos e traços órfãos: Como vimos na seção 4.2, uma das etapas do casamento é a remoção de símbolos e traços órfãos. Porém, esse pós-processamento não garante que todos os traços e símbolos órfãos sejam removidos, assim, uma grande parte dos erros que o casamento apresenta é devido a não remoção desses símbolos e traços. Um exemplo deste problema é apresentado na figura 5.18.

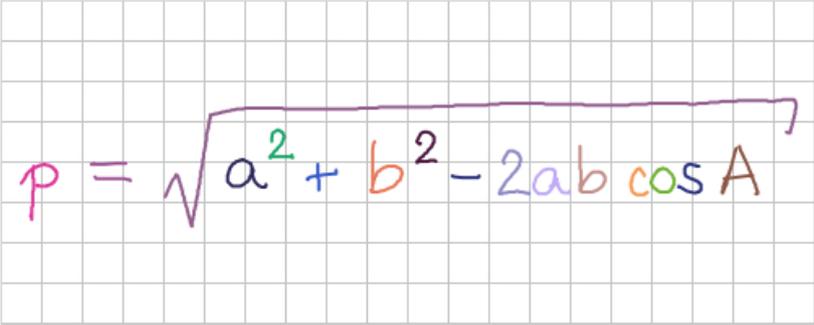
Semelhança de subgrupos: Um problema encontrado em nosso método é a possível

Admin Test Panel

Expression type:	79	User:	Fujita
Expression Status:	Validated	Expression match result:	0.625000

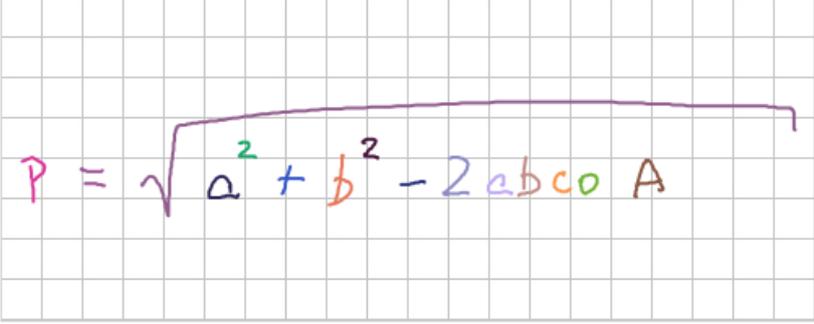
Model Expression

← →



User Transcription

✖ ← →



+ +_2	- -_2	^ 2_4	^ 2_5	^ 2_6	= =_2
A A_2	\sqrt \sqrt_2	a a_3	a a_4	b b_3	b b_4
c c_2	o o_2	p p_2			

Figura 5.15: Caso onde o espaçamento entre os símbolos é desigual. Note que no decorrer da expressão, o espaçamento diminuiu, o que causa um erro no casamento do símbolo 's'.

ambigüidade de classes de símbolos oriunda do fato de um subgrupo de um agrupamento também assemelhar-se a um símbolo. Por exemplo, no caso do símbolo (\ominus), o traço que corta o símbolo pode ser facilmente reconhecido como um sinal de subtração ou um fracionário. Na figura 5.19 apresentamos um exemplo deste erro. No caso, o traço que corta o sinal de adição foi reconhecido como o fracionário pois o custo para o casamento é menor.

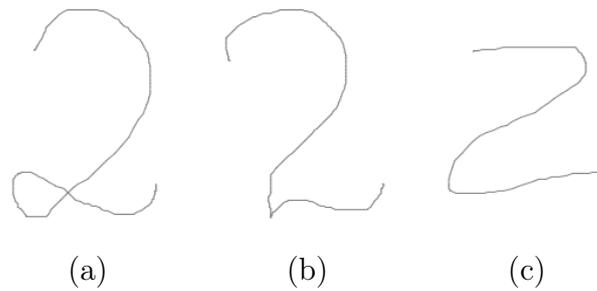


Figura 5.16: Três maneiras distintas (a), (b) e (c) de escrever o símbolo '2'.

5.6 Tecnologias utilizadas

Para a construção do ExpressMath Online foram utilizadas tecnologias baseadas na plataforma Java ¹, versão 1.7, além dos frameworks spring ² e jsf ³, que foram de grande ajuda para o desenvolvimento deste trabalho.

¹<http://www.oracle.com/technetwork/java>

²<http://projects.spring.io/spring-framework/>

³<http://www.oracle.com/technetwork/java/javase/javaserverfaces-139869.html>

Model Expression

$$V_I = \sqrt{D} + \sqrt{D - \epsilon \sqrt{D}}$$

Transcription

$$V_I = \sqrt{D} + \sqrt{D - \epsilon \sqrt{D}}$$

(a)

Model Expression

$$V_I = \sqrt{D} + \sqrt{D - \epsilon \sqrt{D}}$$

Transcription

$$V_I = \sqrt{D} + \sqrt{D - \epsilon \sqrt{D}}$$

(b)

Figura 5.17: Na figura (a) apresentamos a transcrição de uma expressão modelo, e na figura (b) apresentamos o resultado do casamento entre as expressões; o erro de casamento ocorre pois a forma de escrita dos símbolos difere entre as expressões modelo e transcrição.

Admin Test Panel

Expression type: 88 User: Fabricio
 Expression Status: Validated Expression match result: 0.773585

Model Expression

← →

User Transcription

🎨 ← →

■ +_6	■ +_7	■ +_8	■ -_10	■ -_11	■ -_12
■ -_13	■ -_14	■ -_15	■ -_16	■ -_17	■ -_18
■ 2_16	■ 2_17	■ 2_18	■ 2_19	■ 2_20	■ 2_21
■ 2_22	■ 2_23	■ 2_24	■ 2_25	■ 2_26	■ 2_27
■ 2_28	■ 2_29	■ 2_30	■ =_2	■ M_2	■ [_2
■]_2	■ a_10	■ a_11	■ a_12	■ a_13	■ a_14
■ a_8	■ a_9	■ b_10	■ b_11	■ b_12	■ b_13
■ b_14	■ b_8	■ b_9	■ c_10	■ c_11	■ c_12
■ c_13	■ c_14	■ c_8	■ c_9		

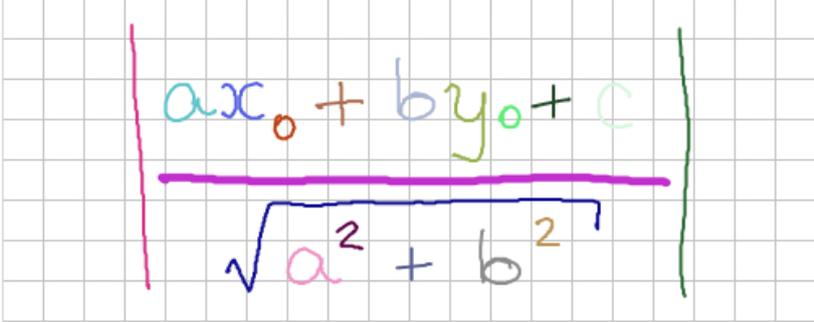
Figura 5.18: Caso de erro de casamento onde restaram traços e símbolos órfãos.

Admin Test Panel

Expression type: 73 User: Alfonso
 Expression Status: Validated Expression match result: 0.944444

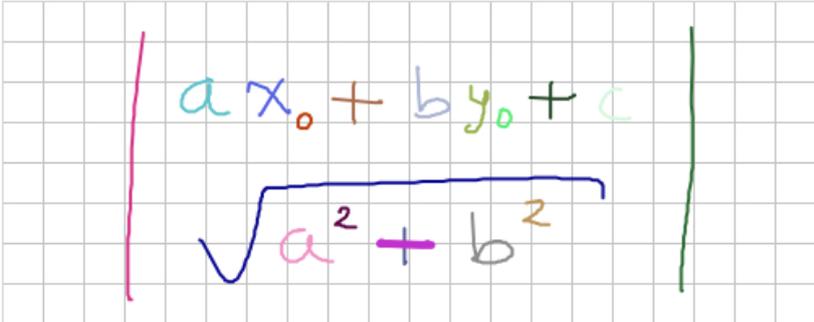
Model Expression

← →



User Transcription

🎨 ← →



+_4	+_5	+_6	0_3	0_4	2_3
2_4	\frac_2	\sqrt_2	a_3	a_4	b_3
b_4	c_2	x_2	y_2	L_3	L_4

Figura 5.19: Erro de casamento devido a semelhança entre o subgrupo do sinal de adição (-) com o símbolo fracionário.

Capítulo 6

Conclusões

Neste trabalho foi proposto um sistema web, o ExpressMath Online, para o tratamento de dados de EMs. A principal característica desse sistema é servir de base para a criação de amostras de EMs anotadas com informação de *ground-truth*, o desenvolvimento de métodos de reconhecimento de expressões matemáticas manuscritas *online*, e a análise de desempenho dos reconhecedores. Para tanto, ele oferece funcionalidades básicas necessárias para a coleta, armazenamento e gerenciamento de dados de EMs. Oferece também funcionalidades para a inclusão de novos métodos de reconhecimento, aplicação automática desses métodos sobre um conjunto pré-selecionado de EMs, e visualização dos resultados.

Junto ao sistema está acoplado um método para casamento de EMs. O casamento de EMs tem por finalidade auxiliar a criação de múltiplas instâncias de uma expressão, devidamente anotadas com informação de *ground-truth*. O método consiste em coletar transcrições de expressões modelo e então automaticamente associar cada símbolo da expressão modelo ao correspondente símbolo da expressão transcrita. Com isso, toda informação de *ground-truth* da expressão modelo pode ser transferida à expressão transcrita. O método desenvolvido neste trabalho não parte do pressuposto de que os símbolos na expressão transcrita estão corretamente segmentados, como ocorre em um trabalho anterior de Honda (2013).

Para testar as funcionalidades disponíveis no sistema, foram realizados experimentos com o método de casamento proposto usando o próprio sistema. Para isso foi utilizado um conjunto de 932 expressões, transcrições de 54 tipos de expressões. Foi obtida uma taxa de acerto de 79,61%. Embora esta taxa ainda esteja distante de um valor ideal, o método de casamento pode ser útil para reduzir o trabalho necessário para gerar amostras de EMs com informação de *ground-truth*. Facilitar essa tarefa é extremamente importante para obter uma estimativa robusta do desempenho de métodos de reconhecimento.

As implementações do sistema descrito neste trabalho, assim como todos os insumos necessários para construir o sistema estão disponíveis em <https://bitbucket.org/davigrangeiro/express-match>.

6.0.1 Trabalhos futuros

Pretendemos em breve divulgar o sistema, suas funcionalidades e características em um artigo e disponibilizar o sistema na web. A este sistema deverá ser acoplado o método de casamento de expressões proposto em [Hirata e Julca-Aguilar \(2015\)](#) (que diferentemente do proposto aqui, parte do pressuposto de que os símbolos estão previamente segmentados nas expressões). Deverá também ser acoplado um método de reconhecimento de expressões, que se encontra em desenvolvimento ([Julca-Aguilar et al., 2015](#)).

A partir dos resultados dos testes realizados no sistema ExpressMath Online, constatamos que certos pontos necessitam ser aprimorados. A seguir são descritas algumas oportunidades e desafios que devem aprimorar a experiência na utilização do sistema e melhorar o desempenho da técnica de casamento de expressões proposta.

Aprimoramento da interface do sistema: Ao decorrer da construção do ExpressMath Online, alguns problemas na navegação ficaram aparentes, devido a tentativa de construção de um sistema baseado em janelas e totalmente touch-friendly. Portanto é necessário estudar soluções para os problemas ou adotar bibliotecas especializadas no tipo de comportamento citado para a realização do mesmo.

Aprimoramento da técnica de casamento de expressões sem restrições: Apesar dos resultados do método de casamento de expressões serem significativos, eles ainda estão longe de serem ótimos para ser adotado amplamente. Portanto é necessário que a técnica seja aprimorada para a obtenção de melhores resultados. Essa melhora agilizaria o processo de criação de banco de dados de EMs anotadas com ground-truth, uma vez que os validadores das expressões teriam menos trabalho na correção do casamento das expressões; permitindo que um maior número de expressões sejam validadas.

Referências Bibliográficas

- Awal et al. (2009)** A.-M. Awal, H. Mouchere e C. Viard-Gaudin. Towards handwritten mathematical expression recognition. *Document Analysis and Recognition, 2009. ICDAR '09. 10th International Conference*. Citado na pág. [1](#)
- Awal et al. (2010)** Ahmad-Montaser Awal, Harold Mouchère e Christian Viard-Gaudin. The problem of handwritten mathematical expression recognition evaluation. *12 th. International conference on frontiers in handwriting recognition*, páginas 646–651. Citado na pág. [10](#)
- Belongie et al. (2002)** S. Belongie, J. Malik e J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Anal. Mach. Intell.*, páginas 509–522. Citado na pág. [12](#)
- Bresler et al. (2012)** M. Bresler, D. Prusa e V. Hlavac. MfrDB: Database of annotated on-line mathematical formulae. *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference*, páginas 542 – 547. Citado na pág. [2](#)
- Chan e Yeung (2000)** Kan-Fai Chan e Dit-Yan Yeung. Mathematical expression recognition: a survey. *International Journal on document analysis and Recognition*, páginas 3–15. Citado na pág. [1](#), [8](#)
- Hirata e Honda (2011)** Nina S. T. Hirata e Willian Y. Honda. Automatic labeling of handwritten mathematical symbols via expression matching. *8th IAPR-TC-15 International Workshop, GbRPR 2011, Münster, Germany, May 18-20. Proceedings*, páginas 295 – 304. Citado na pág. [10](#)
- Hirata e Julca-Aguilar (2015)** Nina S. T. Hirata e Frank D. Julca-Aguilar. Matching based ground-truth annotation for online handwritten mathematical expressions. *Pattern Recognition*, 48(3):837 – 848. ISSN 0031-3203. doi: <http://dx.doi.org/10.1016/j.patcog.2014.09.015>. URL <http://www.sciencedirect.com/science/article/pii/S0031320314003768>. Citado na pág. [1](#), [3](#), [4](#), [10](#), [15](#), [56](#)
- Honda (2013)** Willian Yukio Honda. Rotulação de símbolos matemáticos manuscritos via casamento de expressões. Dissertação de Mestrado, Instituto de Matemática e Estatística, Universidade de São Paulo, Brasil. Citado na pág. [vii](#), [xi](#), [3](#), [4](#), [10](#), [11](#), [14](#), [15](#), [29](#), [55](#)
- Huang e Kechadi (2007)** B. Q. Huang e M-T Kechadi. A structural analysis approach for online handwritten mathematical expressions. *International Journal of Computer Science and Network Security, IJCSNS'07*. Citado na pág. [1](#)
- Huang et al. (2007)** B. Q. Huang, Y. B. Zhang e M-T Kechadi. Preprocessing techniques for online handwriting recognition. Em *Proceedings of the Seventh International Conference*

- on Intelligent Systems Design and Applications*, páginas 793–800. IEEE Computer Society. Citado na pág. 12
- Julca-Aguilar et al. (2014)** F. D. Julca-Aguilar, N. S. T. Hirata, C. Viard-Gaudin, H. Mouchère e S. Medjkoune. Mathematical symbol hypothesis recognition with rejection option. Em *14th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, páginas 500–505. Citado na pág.
- Julca-Aguilar e Hirata. (2012)** Frank D. Julca-Aguilar e Nina S. T. Hirata. Express-match: a system for creating ground-truthed datasets of online mathematical expressions. *Proc. 9th IAPR International Workshop on Document Analysis Systems*. Citado na pág. 2, 15, 37
- Julca-Aguilar et al. (2015)** Frank D. Julca-Aguilar, Harold Mouchère, Christian Viard-Gaudin e Nina S. T. Hirata. Top-down graph grammar parsing for recognition of online handwritten mathematical expressions. Em *To appear in Proceedings of CIARP 2015*. Citado na pág. 56
- Koschinski et al. (1995)** M. Koschinski, H.-J. Winkler e M. Lang. Segmentation and recognition of symbols within handwritten mathematical expressions. *1995 Int. Conference on Acoustics, Speech, and Signal Processing (ICASSP-95)*, 4:2439 – 2442. Citado na pág. 2, 29, 32
- Kuhn (1955)** Harold W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2:83 – 97. Citado na pág. 11
- Labahn et al. (2008)** George Labahn, Edward Lank, Mirette Marzouk, Andrea Bunt, Scott MacLean e David Tausky. Mathbrush: A case study for pen-based interactive mathematics. *EUROGRAPHICS Workshop on Sketch-Based Interfaces and Modeling*. Citado na pág. 2
- Laponte (2008)** Adrien Laponte. Issues in performance evaluation of mathematical notation recognition systems. Dissertação de Mestrado, Queens’s Univerity. Citado na pág. 9
- Matsakis (1999)** Nicholas E. Matsakis. Recognition of handwritten mathematical expressions. Dissertação de Mestrado, Massachusetts Institute of Technology. Citado na pág. 2, 29, 32
- Mouchère et al. (2013)** Harold Mouchère, C. Viard-Gaudin, Richard Zanibbi, U. Garain, D. H. Kim e J. H. Kim. ICDAR 2013 CROHME: Third international competition on recognition of online handwritten mathematical expressions. Citado na pág. xv, 2, 8
- Munkres et al. (1957)** J. Munkres, P. Foggia, C. Sansone e M. Vento. Algorithms for the assignment and transportation problems. *Journal of the Society of Industrial and Applied Mathematics*, 5(1):32 – 38. Citado na pág. 11
- Noma et al. (2011)** Alexandre Noma, Alvaro Pardo e Roberto M. Cesar Jr. Structural matching of 2D electrophoresis gels using deformed graphs. *Pattern Recognition. Lett*, 32: 3 – 11. Citado na pág. 14
- Toumit et al. (1999)** J.-Y. Toumit, S. Garcia-Salicettia e H. Emptoz. A hierarchical and recursive model of mathematical expressions for automatic reading of mathematical documents. *Document Analysis and Recognition, 1999. ICDAR '99. Proceedings of the Fifth International Conference*. Citado na pág. 2

- Winkler e Lang (1997)** Hans-Jürgen Winkler e Manfred Lang. Online symbol segmentation and recognition in handwritten mathematical expressions. Em *In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, páginas 3377–3380. Citado na pág. [29](#)
- Xie (2007)** Xiaofang Xie. *On the recognition of handwritten mathematical symbols*. Tese de Doutorado, The University of Western Ontario - London. Citado na pág. [2](#)