



MÉTRICA PARA AVALIAÇÃO DO VALOR SEMÂNTICO DE UM SIGNO ENTRE/EM SISTEMAS COM ACOPLAMENTO PARCIAL

Trabalho de Conclusão de Curso

Engenharia da Computação

Larissa Santiago de Lima Soares
Orientador: Prof. Fernando Buarque de Lima Neto



**UNIVERSIDADE
DE PERNAMBUCO**

**Universidade de Pernambuco
Escola Politécnica de Pernambuco
Graduação em Engenharia de Computação**

LARISSA SANTIAGO DE LIMA SOARES

**MÉTRICA PARA AVALIAÇÃO DO
VALOR SEMÂNTICO DE UM SIGNO
ENTRE/EM SISTEMAS COM
ACOPLAMENTO PARCIAL**

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia de Computação pela Escola Politécnica de Pernambuco – Universidade de Pernambuco.

Recife, Novembro de 2013.

De acordo

Recife

____/____/____

Fernando Buarque de Lima Neto

“Daria tudo que sei pela metade do que ignoro”
Atribuído a René Descartes

Dedicado à minha família.

Agradecimentos

Agradeço a Deus pela minha vida e pela oportunidade de fazer parte do curso de Engenharia da Computação da Universidade de Pernambuco, onde consegui completar cinco anos de graduação com a certeza que estes foram anos de evolução pessoal e profissional. Agradeço a todos os presentes desde quando eu não compreendia o significado de “Universidade” e todos àqueles que deram sentido à palavra.

Aos meus queridos mãe, pai e irmão agradeço pela paciência e amor despendidos em todos os momentos. Sem o apoio de vocês eu não concluiria tão bem mais esta etapa. Agradeço também a todos os funcionários da universidade. E aos professores pela dedicação e perseverança numa profissão tão importante; em especial ao meu orientador pela atenção desde os primeiros períodos acadêmicos e por inquietar as mentes dos alunos para alcançar horizontes mais distantes.

Agradeço a todos os amigos que fiz durante esses anos, que dividiram o sono comigo todas as manhãs e tornaram as aulas mais divertidas. Vocês, agora parceiros de profissão, serão especiais independente do passar dos anos. Sou feliz por ter construído amizades sinceras ao longo da graduação. Aos membros do time de semiótica computacional e a todos os demais que de alguma forma me ajudaram a construir este trabalho destino um agradecimento carinhoso pelos ótimos momentos compartilhados.

Por fim, agradeço pela felicidade de afirmar que conclui a graduação em Engenharia da Computação e por estar preparada para as próximas etapas e futuros desafios.

Larissa Santiago de Lima Soares

Resumo

A semiótica é uma ciência que estuda os signos e os processos mentais responsáveis pela significação. Um signo representa algo para alguém. Notadamente, a semiótica é transdisciplinar e analisa-la é fundir as mais diversas ciências de conhecimentos biológicos, sociais e lógicos na busca pelo sentido das coisas que rodeiam o ser humano. Os pensamentos são únicos e particulares para cada indivíduo, conseqüentemente, o processo semiótico que dá sentido a um signo é diferente na mente de cada ser humano; depende do contexto a que o signo está inserido e das experiências mentais já vividas pelo indivíduo. Servir-se de técnicas de inteligência artificial para compreender fenômenos naturais da mente humana é uma tarefa interessante e desafiadora. Este trabalho propõe uma combinação da semiótica com os mapas auto-organizáveis de Kohonen (também chamados de redes neurais artificiais de Kohonen) a fim de desenvolver uma métrica que auxilie a avaliação do valor semântico de um signo para um sistema e entre sistemas, por exemplo, a avaliação do significado de um signo para um ser humano e a diferença entre os significados atribuídos a um signo por indivíduos distintos em um processo de comunicação. A construção da métrica possibilitou objetivar a distância conceitual de signos subjetivos através de medidas de distâncias matemáticas. Desta forma, é possível comparar numericamente o entendimento de dois sistemas sobre um dado conceito. Medir o valor semântico de um signo para um sistema é o passo inicial rumo ao aperfeiçoamento da fidelidade da comunicação humano-humano e até humano-máquina, visto que pode diminuir a dissimilaridade entre significados e proporcionar menor falha de comunicação. A métrica, usada como componente de sistemas maiores, poderá ser útil também para comparar conceitos, aproveitar-se da comparação para construir *feedbacks* mais apropriados e até recuperar conceitos subjetivos baseados em experiências anteriores com uma significação que pode ser recontextualizada.

Abstract

Semiotics is a science that studies signs and mental processes accountable for signification. A sign represents something to someone. Notably, semiotics is transdisciplinary, thus it can blend concepts of most varied sciences of biological, social and logical knowledge on the search for the meaning in the things that surround humans. Thoughts are unique and particular for each individual, hence the semiotic process that gives meaning to a sign happens differently on the mind of every human being; depends on the context to which the sign is inserted and depends on the mental experiences already lived by the individual. Make use of artificial intelligence techniques to comprehend natural phenomenon of the human mind is an interesting and challenging task. This work proposes a combination among semiotics and Kohonen's self-organizing maps (also known as Kohonen's artificial neural network) in order to develop a metric to assist the evaluation of the semantic value of a sign for a system and between systems, that is, the evaluation of the meaning of a sign for a human being and the difference between the meanings given to a sign by distinguished individuals in a communication process. The construction of the metric allowed to objectify the conceptual distance of subjective signs through mathematical distances. Thereby, it is possible to compare numerically the understanding on a given concept by two systems. Measuring the semantic value of a sign for a system is the initial step towards improving the fidelity of communication human-human and human-machine, since it can decrease the dissimilarity between meanings and provide lower communication failure. The proposed metric used as a component of larger systems may be also helpful to compare concepts, to take advantage of comparison to offer more appropriate feedbacks and retrieve subjective concepts based on past experiences with a meaning that can be recontextualized.

Sumário

Resumo.....	vii
Abstract.....	viii
Índice de Figuras.....	xi
Índice de Tabelas.....	xii
Lista de Símbolos e Siglas.....	xiii
Capítulo 1 Introdução	1
1.1 Motivação e Caracterização do Problema	1
1.2 Objetivos e Metas	2
1.3 Estrutura do Documento	3
Capítulo 2 Fundamentação Teórica	4
2.1 Semiótica	4
2.1.1 Signo	5
2.1.2 As Classificações Triádicas de Peirce.....	7
2.1.3 As Dimensões da Semiose de Morris	8
2.2 Comunicação.....	9
2.2.1 Fidelidade da Comunicação.....	11
2.3 Métricas e Medidas.....	12
2.3.1 Medidas de Similaridade/Dissimilaridade.....	12
2.4 Mapa Auto-organizável de Kohonen.....	14
2.4.1 Redes Neurais Artificiais	14
2.4.2 Algoritmos de <i>Clusterização</i>	15
2.4.3 Self-Organizing Maps (SOM)	17
2.4.4 <i>K-means versus</i> SOM	20
Capítulo 3 Proposição da Métrica	21
3.1 Visão Geral do Modelo	21
3.2 Representação Vetorial do Conceito	24
3.3 Comparação e Interpretação	26
Capítulo 4 Experimentos e Resultados	28
4.1 Cenário Um: Suporte à Decisão	28
4.1.1 Representação da Decisão de Compra	29
4.1.2 Treinamento e Teste	31

4.2 Cenário Dois: Educação	34
4.2.1 Cenário e Base de Dados	35
Capítulo 5 Conclusões e Trabalhos Futuros	36
5.1 Conclusões	36
5.2 Trabalhos Futuros.....	37
Bibliografia.....	38
Anexo A Decisões apropriadas de decisores e subordinados.....	42
Apêndice A Vetores numéricos das decisões	44

Índice de Figuras

Figura 1. Modelo diádico de Saussure para o signo.....	5
Figura 2. Modelo triádico de Peirce para o signo.	6
Figura 3. Modelo do processo de comunicação de Shannon e Weaver.....	10
Figura 4. Modelo de comunicação de Jakobson	10
Figura 5. Distâncias Manhattan (Dm) e Euclidiana (DE)	14
Figura 6. Exemplo de taxonomia para técnicas de <i>clusterização</i>	16
Figura 7. Estrutura de um mapa auto-organizável de Kohonen.	17
Figura 8. Representação dos vetores de dados e dos neurônios no SOM.	18
Figura 9. Vizinhança topológica de um neurônio.....	19
Figura 10. Exemplo de comunicação com semiose.	22
Figura 11. Modelo da métrica proposta.	24
Figura 12. Vetor de conceitos objetivos e subjetivos.	25
Figura 13. Camisa.	26
Figura 14. SOM gerado pela ferramenta Spice-SOM.....	33

Índice de Tabelas

Tabela 1. Relações entre categorias e tricotomias de Peirce.	8
Tabela 2. Atributos de uma camisa.....	25
Tabela 3. Vetor representativo de uma camisa.....	26
Tabela 4. Atributos subjetivos de uma decisão de compra.....	29
Tabela 5. Normalização dos atributos.....	30
Tabela 6. Valores normalizados para as prioridades	31
Tabela 7. Resposta do decisor (D) e dois subordinados (S1 e S2)	31
Tabela 8. Representação do vetor do decisor (D1)	31

Lista de Símbolos e Siglas

2D – Duas dimensões.

ANN – Do inglês, *Artificial Neural Networks*.

BIPM – Do francês, *Bureau International des Poids et Mesures*.

BMU – Do inglês, *Best Matching Unit*.

EaD – Ensino à Distância

IEEE – Do inglês, *Institute of Electrical and Electronics Engineers*.

ISO – Do inglês, *International Organization for Standardization*

ITSs – Do inglês, *Intelligent Tutoring Systems*

JML – Do inglês, *Java Machine Learning*

RGB – Do inglês, *Red, Green and Blue*.

SI – Do francês, *Le Système International D'unités*.

SOM – Do inglês, *Self-Organizing Maps*

Capítulo 1

Introdução

Neste capítulo são explicados os motivos que deram origem a esta pesquisa seguido pelas características do problema explorado. Estão descritos o objetivo principal e os objetivos específicos do trabalho e por fim há uma breve descrição da estrutura do documento e dos focos dos capítulos posteriores.

1.1 Motivação e Caracterização do Problema

O cérebro humano é um computador natural muito poderoso, onde o processamento de informações ocorre em um alto grau de complexidade através do que costumamos chamar de mente. Ao interagir com o mundo cada ser humano interpreta fatos, objetos e sentimentos de uma forma particular, ou seja, dentro do universo semântico de conexões criado em sua própria mente.

A semiótica [1] estuda como o ser humano, um sistema natural, significa aquilo que o rodeia, isto é, estuda os signos [2]. Devido a isso é um campo de pesquisa vasto e de grande importância antropológica. Através da semiótica é possível analisar o valor semântico [3] que o ser humano confere aos signos ou conjunto de signos que percebe.

O processo semiótico que acontece na mente de cada ser humano é único [1][4], desta forma, quando dois seres interagem em um processo de comunicação não é possível afirmar que o entendimento mútuo foi alcançado simplesmente pelo fato de a mensagem ter sido transmitida, pois o entendimento da mesma cabe ao universo mental daquele que a recebeu.

A sociedade estabeleceu padrões objetivos para facilitar a comunicação, como letras, números, pesos, medidas de tamanho e distância, etc. Porém, ao se tratar de conceitos subjetivos não há padrões. Por exemplo, quando o diretor de uma empresa diz ao seu gerente para comprar a máquina de café mais barata,

provavelmente o gerente corresponderá à expectativa do diretor. Porém, se o diretor pede pela máquina de café com o design mais bonito, o gerente poderá não corresponder como deveria, já que sua concepção de beleza pode ser diferente.

Nesse contexto, uma métrica para avaliar o valor semântico de um signo ou conjunto de signos para um sistema e a conseqüente comparação com o valor para outro sistema é importante para medir a diferença do entendimento de um conceito, aperfeiçoar a comunicação e até auxiliar no desenvolvimento de ideias levando em consideração conceitos subjetivos.

1.2 Objetivos e Metas

O objetivo central deste trabalho é definir uma métrica para auxiliar a avaliação semântica de um signo para um sistema e quão diferente se torna a representação deste signo em relação a outro sistema. A métrica será baseada em semiótica e se utilizará dos mapas auto-organizáveis de Kohonen [5]. Para alcançar o resultado esperado foram traçados os seguintes objetivos específicos:

- Realizar um estudo sobre semiótica e comunicação;
- Realizar um estudo sobre os mapas auto-organizáveis de Kohonen;
- Realizar um estudo sobre métricas, escalas e medidas de similaridade/dissimilaridade;
- Desenvolver e formalizar a nova métrica e todos os detalhes que a compõem;
- Estudar dois domínios de teste para a métrica: suporte à decisão e educação;
- Realizar ao menos um experimento com a métrica para cada um dos dois domínios anteriores, totalizando dois cenários de teste;
- Analisar os resultados e contribuições da métrica.

1.3 Estrutura do Documento

Este documento está dividido em cinco capítulos, cujos resumos estão descritos abaixo:

- **Capítulo 1: Introdução**

Este capítulo contém os textos que introduzem este trabalho, tais quais a motivação, a caracterização do problema e os objetivos gerais e específicos.

- **Capítulo 2: Fundamentação Teórica**

Este capítulo agrupa os principais conceitos para fundamentar a teoria contida no trabalho. Logo, serão descritas as principais informações sobre Semiótica, comunicação, métricas e medidas e mapas auto-organizáveis de Kohonen.

- **Capítulo 3: Métrica Proposta**

Neste capítulo é apresentada a principal contribuição do trabalho, a métrica para avaliação do valor semântico entre/em sistemas com acoplamento parcial; a métrica reúne os conceitos explicados no capítulo anterior.

- **Capítulo 4: Experimentos e Resultados**

Este capítulo apresenta dois cenários de domínios distintos, úteis para os testes com a métrica proposta. O primeiro cenário envolve suporte à decisão e o segundo educação.

- **Capítulo 5: Conclusões e Trabalhos Futuros**

Este capítulo contém os textos que concluem o documento, destacando as contribuições da proposta e enunciando os possíveis trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Este capítulo tem por objetivo embasar a métrica proposta neste trabalho e apresentar as informações básicas necessárias ao entendimento do documento. As duas primeiras seções tratam de dois conceitos abstratos intrinsecamente conectados, a semiótica e a comunicação. A primeira contém uma introdução ao conceito de semiótica, uma análise temporal e a apresentação de conceitos encontrados na literatura de Charles S. Peirce [1] e Charles W. Morris [3]. A segunda explora as principais características do processo de comunicação e a busca pela qualidade do mesmo. A terceira seção, que aborda métricas e medidas de similaridade e dissimilaridade, auxilia a compreensão da última seção, cujo foco principal é o mapa auto-organizável de Kohonen.

2.1 Semiótica

A teoria dos signos é estudada desde o período da Grécia antiga, através de filósofos como Platão, Aristóteles e Aurélio Agostinho, passando pelos séculos XVII e XVIII com Thomas Hobbes e John Locke, chegando aos séculos XIX e XX com o matemático e filósofo Charles Sanders Peirce [6] e o linguista Ferdinand de Saussure [7] e se firmando como uma ciência dos signos, da significação e da cultura [8][9].

Locke utilizou o termo grego semiótica para definir a doutrina dos signos [10], assim como Peirce [1], algumas décadas depois. Saussure, contemporâneo a Peirce, analisou a semiótica de uma perspectiva linguística definindo um modelo diádico (composto por duas partes) para o signo [11], enquanto Peirce formalizou a semiótica de um ponto de vista lógico e desenvolveu um modelo triádico para o signo.

No panorama atual, analisando a semiótica de maneira vasta, Lúcia Santaella, pesquisadora brasileira, afirmou que a semiótica é a ciência que tem por

objeto de investigação todas as linguagens possíveis, ou seja, que tem por objetivo o exame dos modos de constituição de todo e qualquer fenômeno como fenômeno de produção de significação e de sentido [4]. Charles W. Morris vai além, quando sustenta a ideia de que mais do que uma ciência, a semiótica é um instrumento das ciências, na medida em que cada ciência faz uso e exprime os seus resultados em termos de signos [3].

Fica claro através da literatura envolvida com a semiótica que o signo é algo inerente à definição formal desta ciência, portanto o entendimento do que vem a ser um signo é essencial para a fundamentação da semiótica. A próxima seção apresentará os modelos diádico e triádico de um signo dos pontos de vista de Saussure e Peirce, enfatizando a formalização proposta por este.

2.1.1 Signo

Saussure, em sua obra *Curso de linguística geral* [11] define signo como a combinação de duas partes: o conceito e o padrão de som; ou através da representação de Daniel Chandler [12], respectivamente significado e significante, como mostra a Figura 1.

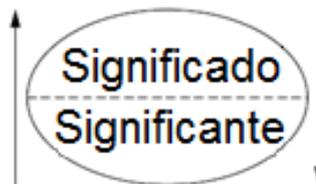


Figura 1. Modelo diádico de Saussure para o signo [12].

O padrão de som (significante) não é realmente um som, já que o som é algo físico, mas é uma impressão psíquica desse som. Hoje em dia, o significante é considerado a parte material e o significado a parte abstrata do signo. A relação entre as duas partes é chamada de significação, porém este laço é arbitrário, como explicou Saussure através da linguística, lembrando que não há uma conexão direta e natural entre o som de uma palavra e o conceito a que se refere. Saussure entendia a semiótica principalmente de um ponto de vista linguista, por isso, neste trabalho, tomaremos o ponto de vista lógico de Peirce para o signo.

Para Peirce, um signo, ou *representamen*, é aquilo que, sob certo aspecto, representa algo para alguém. Dirige-se a alguém, isto é, cria na mente dessa pessoa, um signo equivalente [1].

Uma representação tríadica do signo já havia sido proposta por alguns nomes durante o curso da história [2], mas Peirce se destacou por dedicar sua vida a estudar e escrever uma filosofia científica da linguagem [4] e elaborar uma importante classificação dos signos. Peirce definiu um signo como a combinação de três elementos; objeto, *representamen* ou signo e interpretante:

- Objeto: é aquilo a que o signo se refere.
- *Representamen* ou signo: é a representação imediata do objeto na mente, ou seja, aquilo que funciona como um signo não necessariamente em uma forma material, pois não representa o objeto em todos os seus aspectos, mas é uma referência idealizada do que é objeto.
- Interpretante: é o signo assim criado a partir do primeiro signo [1]; o interpretante não é o intérprete, mas o processo relacional que se cria na mente do intérprete conforme a percepção do objeto [4].

Os três elementos da tríade de Peirce são usualmente ilustrados pela Figura 2, onde a linha tracejada indica que não há uma relação direta entre o objeto e o *representamen* sem que haja o interpretante.

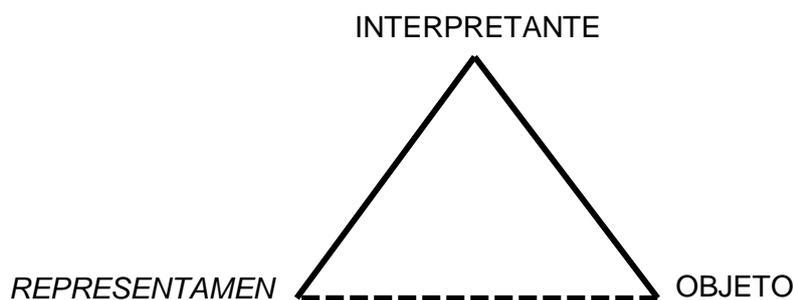


Figura 2. Modelo tríadico de Peirce para o signo [13].

Peirce utiliza a palavra “signo” ora em um sentido mais amplo, onde o signo designa a relação tríadica entre objeto, *representamen* e interpretante e ora em um sentido mais restrito, onde o signo é um elemento da tríade, o *representamen*.

Resumidamente, o signo é um relacionamento do que é representado (o objeto), de como ele é representado (o *representamen*) e de como ele é interpretado (o interpretante) [12].

A partir da relação de representação que o signo mantém com seu objeto, produz-se na mente interpretadora um outro signo que traduz o significado do primeiro (é o interpretante do primeiro). Portanto, o significado de um signo é outro signo [4].

Este processo de cooperação entre os três elementos para a formação do signo é chamado de semiose. A semiose é um processo infinito de significação [14][1], já que cada signo leva a outro signo, que leva a outro signo, e assim sucessivamente. E é um processo que ocorre de maneira única em cada indivíduo, visto que duas pessoas possuindo a mesma estrutura semiótica e percebendo o mesmo objeto, sofreriam a intervenção de fatores como ângulo de visão e a iluminação, provocando um processo semiótico distinto [15].

2.1.2 As Classificações Triádicas de Peirce

Levando em consideração a consciência como um lugar em que os pensamentos interagem e manifestações psicológicas acontecem, Peirce definiu três modos como os fenômenos aparecem à consciência ou três categorias de pensamento: Primeiridade, que é o sentimento imediato e presente das coisas, ou seja, um pensamento primitivo e simples. É o modo de ser sem referência a qualquer outra coisa; Secundidade, que é o modo de ser relacionado a um segundo; e Terceiridade, que é a categoria do pensamento em signos e da interpretação do mundo, possui natureza mais complexa e aproxima as duas outras categorias numa síntese intelectual. De outra forma as três categorias poderiam ser entendidas, respectivamente, por: um *feeling* imediato ou o caráter de apresentação do signo; um sentido de reação ou o caráter de representação do signo; e uma formação de hábito ou o poder interpretativo do signo [1][4][12][16].

Peirce estabeleceu em seus escritos dez tricotomias para classificar os signos, cuja combinação resulta em aproximadamente 64 classes de signos [4]. As tricotomias são relações triádicas entre um fenômeno da primeiridade, outro da

secundidade e o último da Terceiridade. Peirce não chegou a explorar todos os tipos de signos, mas elaborou as dez tricotomias, dentre as quais três são as mais conhecidas e são apresentadas na Tabela 1.

A primeira tricotomia relaciona o *representamen* com ele próprio, ou seja, a característica do signo como signo; desta maneira um signo pode ser uma qualidade (qualissigno), uma coisa ou evento existente e real (sinsigno) ou uma lei, normalmente estabelecida pelos homens (legissigno). A segunda tricotomia relaciona o signo com seu objeto; assim um signo assemelha-se ou imita o objeto (ícone), indica um objeto porque mantém uma conexão direta com ele (índice) ou relaciona-se arbitrariamente com o objeto através de uma convenção social (símbolo). Por fim, a terceira tricotomia relaciona o signo com seu interpretante, desta forma um signo pode não ser verdadeiro nem falso, ser simplesmente um nome (rema), ser uma proposição verdadeira ou falsa (dicisigno) ou participar de um discurso racional com premissas que tendem a veracidade (argumento) [1][8][16].

Tabela 1. Relações entre categorias e tricotomias de Peirce.

Tricotomias Categorias	R-R	R-O	R-I
Primeiridade	qualissigno	ícone	rema
Secundidade	sinsigno	índice	dicisigno
Terceiridade	legissigno	símbolo	argumento

R é *representamen*, O é objeto e I é interpretante.

2.1.3 As Dimensões da Semiose de Morris

Morris definiu a semiose como o processo em que algo funciona como um signo. E a analisa como a combinação de três fatores: o veículo sígnico, o *designatum* e o interpretante, análogos ao *representamen*, o objeto o interpretante definidos por Peirce.

Semântica, pragmática e sintática foram as nomenclaturas estabelecidas por Morris para as relações diádicas presentes na tríade semiótica; respectivamente, são a relação entre signos e objetos, entre signos e intérpretes e a relação dos signos uns com os outros [3]. Semântica é uma relação entre os signos e seus respectivos significados, logo sua principal investigação é a parcela da realidade representada pelo signo. A pragmática estuda o valor de um signo para aqueles que o utilizam, ou seja, a significação social de um signo, que é dependente do contexto e por isso vai além da visão semântica. A sintática estuda a organização de signos simples em signos complexos e suas implicações na semântica e na pragmática.

A partir desses conceitos definidos por Morris é possível que um signo ou um conjunto de signos, semanticamente, possua vários significados e pragmaticamente provoque efeitos interpretativos distintos em intérpretes distintos ou até no mesmo intérprete. E ainda, um signo pode variar sua semântica para um intérprete de acordo com o contexto atual da situação ao qual este intérprete está inserido.

Se considerarmos que um objeto percebido por um sistema foi enviado por outro sistema a fim de transmitir uma informação com um dado valor semântico estamos falando de comunicação entre estes sistemas.

2.2 Comunicação

A palavra comunicação deriva da palavra latina *communicatione*, que significa participar, pôr em comum. Portanto, comunicar é, etimologicamente, tornar alguma coisa comum entre seres, seja essa coisa uma informação, uma sensação, uma experiência, [17][18] etc. Mesmo tendo Franke Dance catalogado diversas definições de comunicação [19], essa conceituação não é simples, visto que a comunicação é diversa e multidisciplinar. A comunicação vista como um processo evolui no tempo e é dinâmica, por isso analisar o processo comunicativo também não é uma tarefa simples, ainda assim, vários autores desenvolveram modelos com o objetivo de simular o funcionamento deste processo [20], isolando elementos e relacionamentos para facilitar a visualização e o entendimento. Apenas dois modelos serão descritos

abaixo, o que é suficiente neste momento para entender elementos básicos que serão úteis ao trabalho.

O modelo da comunicação proposto por Shannon e Weaver, ilustrado na Figura 3, é amplamente aceito no meio matemático/eletrônico. Consiste em um processo linear simples, onde uma fonte de informação elabora e envia uma mensagem, que é convertida em um sinal por um transmissor. Para chegar ao receptor, o sinal passa por um canal (meio de transmissão). Já o receptor, realiza o inverso do transmissor, reconstruindo a mensagem a partir do sinal. Ao passar pelo canal o sinal pode sofrer interferências, alterando seu estado inicial.

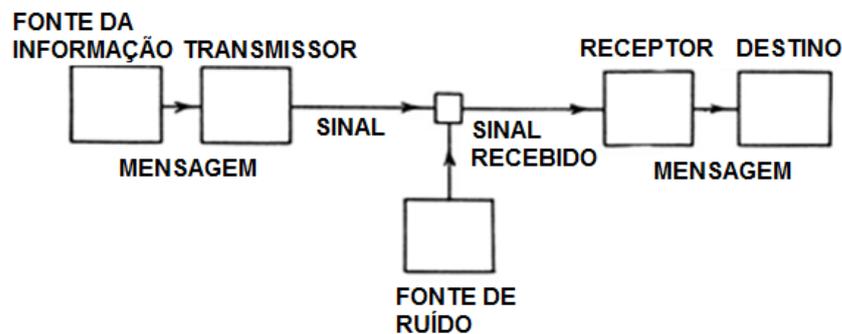


Figura 3. Modelo do processo de comunicação de Shannon e Weaver [21].

Roman Jakobson construiu um modelo do processo de comunicação baseado na linguística [22], fazendo uma ponte com a semiótica. O modelo é ilustrado pela Figura 4. Neste modelo linear existe um destinador e um destinatário para uma mensagem, que por sua vez, possui um contexto. O contato é, além do canal físico de transmissão, a conexão psicológica entre destinador e destinatário, que compartilham um código de entendimento.

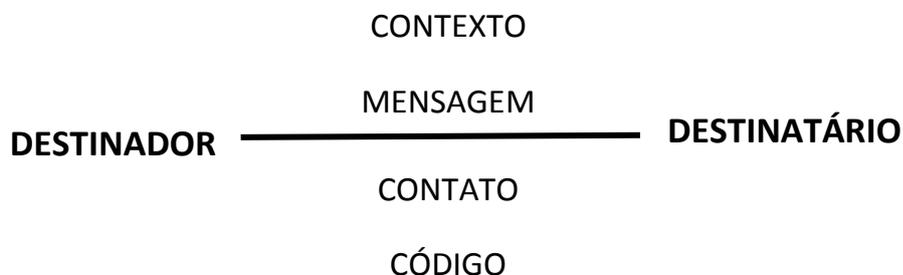


Figura 4. Modelo de comunicação de Jakobson [22].

Estes modelos, apesar de inicialmente inseridos em áreas de conhecimento distintas, possuem elementos básicos em comum. Alguns destes elementos serão utilizados no trabalho daqui em diante, como o emissor, o receptor, a mensagem ou objeto de comunicação, o contexto e o código.

Como visto no modelo de Shannon e Weaver, o sinal pode sofrer interferências. Ao explorar essa ideia os autores falavam de interferências do ponto de vista eletrônico, ou seja, falhas de transmissão que diminuía a qualidade do sinal, isto é reduzia a fidelidade da comunicação.

2.2.1 Fidelidade da Comunicação

Havendo um objetivo a comunicar e uma resposta a obter, um comunicador espera que a sua comunicação seja a mais fiel possível e por fidelidade entende-se que ele obterá o que quer [23]. Apesar do processo linear de comunicação “terminar” quando o objeto é recebido pelo sistema receptor, o propósito maior da comunicação não é apenas transmitir, é receber exatamente o que se desejou enviar. Desta forma, a comunicação entre sistemas é aperfeiçoada à medida que o entendimento do sistema receptor sobre o objeto recebido aproxima-se de cem por cento.

Quando se trata de comunicação entre sistemas digitais existem métodos que detectam e/ou corrigem erros no objeto recebido, tais como o código de Hamming [24] e o *checksum* [25], porém, para sistemas naturais como os seres humanos não existem códigos padronizados para detectar ou corrigir o erro na comunicação, pois as interferências são de ordem mais complexa.

Do ponto de vista semiótico, as falhas de comunicação podem se originar no momento em que o emissor codifica a informação, pois ele pode não reproduzir bem o que está pensando, assim como no momento em que o receptor decodifica a informação e lhe dá um significado, pois este pode ser diferente do significado originado pelo emissor.

2.3 Métricas e Medidas

Os nomes métrica e medida têm alguma sobreposição. Partindo das definições matemáticas encontradas em dicionários e glossários, medida é uma quantidade fixa que serve para avaliar extensões ou quantidades mensuráveis. Métrica é definida pelo IEEE (do inglês, *Institute of Electrical and Electronics Engineers*) como uma medida quantitativa do grau em que um sistema, componente, ou processo possui um dado atributo [26].

Métrica e medida são conceitos semelhantes, mas em geral, utilizamos a palavra medida para atributos concretos ou objetivos e a palavra métrica para atributos de cunho mais subjetivo.

2.3.1 Medidas de Similaridade/Dissimilaridade

A similaridade e a dissimilaridade são medidas opostas de avaliação de distância, a similaridade avalia quão próximos estão dois elementos ou a semelhança entre eles; a dissimilaridade avalia quão distante estão dois elementos ou a diferença entre eles. Quanto maior é o coeficiente de similaridade, mais próximos estão os elementos e quanto maior é a distância, mais diferença há entre os elementos.

Existem diversas medidas de similaridade e dissimilaridade na literatura, tais como as distâncias euclidiana, Manhattan, Minkowski, Jaccard, correlação de Pearson, Chebyshev, Hamming [27][28][29], etc. Para ser considerada uma medida de distância é necessário obedecer às regras abaixo.

Supondo três elementos $\bar{x} = (x_1, x_2 \dots x_p)$, $\bar{y} = (y_1, y_2 \dots y_N)$ e $\bar{z} = (z_1, z_2 \dots z_N)$ de dimensões N em um sistema de coordenadas cartesiano e a distância d entre dois deles:

- Regra 1: $d(\bar{x}, \bar{y}) \geq 0$ onde $d(\bar{x}, \bar{y}) = 0$ se $\bar{x} = \bar{y}$;
- Regra 2: $d(\bar{x}, \bar{y}) = d(\bar{y}, \bar{x})$;
- Regra 3: $d(\bar{x}, \bar{y}) + d(\bar{y}, \bar{z}) \geq d(\bar{x}, \bar{z})$.

Para este trabalho iremos detalhar a distância Minkowski que é uma generalização das distâncias Manhattan (também chamada de distância *city block*), e Euclidiana [27][28][29]. Minkowski é uma distância entre dois pontos e será útil ao entendimento da próxima seção. A distância Minkowski entre os pontos \bar{x} e \bar{y} é indicada pela equação 1.

$$M(\bar{x}, \bar{y}) = \left(\sum_{k=1}^n (x_k - y_k)^p \right)^{\frac{1}{p}} \quad (1)$$

Como foi mencionado, a distância de Minkowski é uma generalização. Normalmente ela é utilizada com os valores de p igual a 1 ou 2. Quando $p=1$ a medida chama-se distância de Manhattan, indicada pela equação 2. Lê-se a equação como a soma das diferenças absolutas das coordenadas de dois elementos.

$$Dm(\bar{x}, \bar{y}) = \sum_{k=1}^n |x_k - y_k| \quad (2)$$

Quando $p=2$ chama-se distância Euclidiana [27]. A distância euclidiana é simples e bastante utilizada. A distância pode ser entendida através da equação 3 como a raiz quadrada da soma dos quadrados das diferenças entre as coordenadas cartesianas dos dois elementos.

$$DE(\bar{x}, \bar{y}) = \sqrt{\left(\sum_{k=1}^n (x_k - y_k)^2 \right)} \quad (3)$$

A diferença entre as distâncias Manhattan e Euclidiana pode ser vista na Figura 5.



Figura 5. Distâncias Manhattan (Dm) e Euclidiana (DE)

2.4 Mapa Auto-organizável de Kohonen

O mapa auto-organizável de Kohonen [5], ou mapa auto-organizável ou simplesmente SOM (do inglês *Self-organizing map*) é um tipo de rede neural artificial desenvolvida pelo professor Teuvo Kohonen [30] em 1982. É uma rede treinada sem supervisão e com a capacidade transformar dados multidimensionais em um mapa, geralmente composto por duas dimensões (2D).

Para compreender um SOM é essencial entender o que é uma rede neural artificial.

2.4.1 Redes Neurais Artificiais

As redes neurais artificiais ou ANNs (do inglês *Artificial Neural Networks*) são modelos que têm como metáfora o funcionamento do cérebro humano, que é um computador altamente complexo, não linear e paralelo. De maneira geral, uma ANN é um sistema constituído por elementos de processamento interconectados, chamados neurônios, os quais estão dispostos em camadas (uma camada de entrada, uma ou várias intermediárias e uma de saída) [31]. O conhecimento é adquirido pela rede através de um processo de aprendizagem cuja função é modificar os pesos da rede; os pesos são forças de conexão entre os neurônios e são utilizados para armazenar o conhecimento adquirido [32]. A aprendizagem de uma ANN pode ser de dois tipos [31][32]:

- Aprendizagem supervisionada ou com professor: as saídas desejadas são fornecidas à rede, que será treinada com influência desse conhecimento

prévio, ajustando os pesos de acordo com a diferença entre o resultado gerado e o desejado.

- Aprendizagem não supervisionada ou auto-organizada: as saídas desejadas não são fornecidas à rede, que será treinada apenas com os dados de entrada. Conseqüentemente, o resultado é fruto das regras de alocação geradas internamente através das características dos dados de entrada.

2.4.2 Algoritmos de *Clusterização*

A *clusterização* ou a capacidade de criar *clusters* é uma das importantes temáticas na área de aprendizagem não supervisionada. *Clusterização* é um processo que organiza objetos em grupos onde os elementos são de alguma forma semelhantes.

Na literatura existem classificações para os diversos algoritmos já desenvolvidos para realizar *clusterização* de dados [28][29]. Os algoritmos podem ser classificados basicamente em hierárquicos e particionais, estes dois tipos ainda podem ser subdivididos em diversas outras classificações. Os algoritmos hierárquicos formam *clusters* gradativamente através de aglomerações ou divisões, gerando uma hierarquia de *clusters*. Já os algoritmos particionais dividem o conjunto de elementos em k subconjuntos, realizando um processo de migração de elementos entre os subconjuntos até que a configuração ideal seja encontrada.

Não será necessário para este trabalho entrar em detalhes sobre a classificação dos algoritmos, mas a Figura 6 mostra um exemplo de taxonomia proposta no artigo “*Data Clustering: A Review*” [33], baseado no livro “*Algorithms for Clustering Data*” [34], que possui um autor em comum com o artigo.

Outras abordagens são possíveis, assim como classificações mais específicas, tal como algoritmos de *clusterização* baseados em lógica Fuzzy [35] ou baseados em estratégias evolucionárias [36]. Porém, neste momento o foco será dado a apenas um algoritmo, o *k-means*.

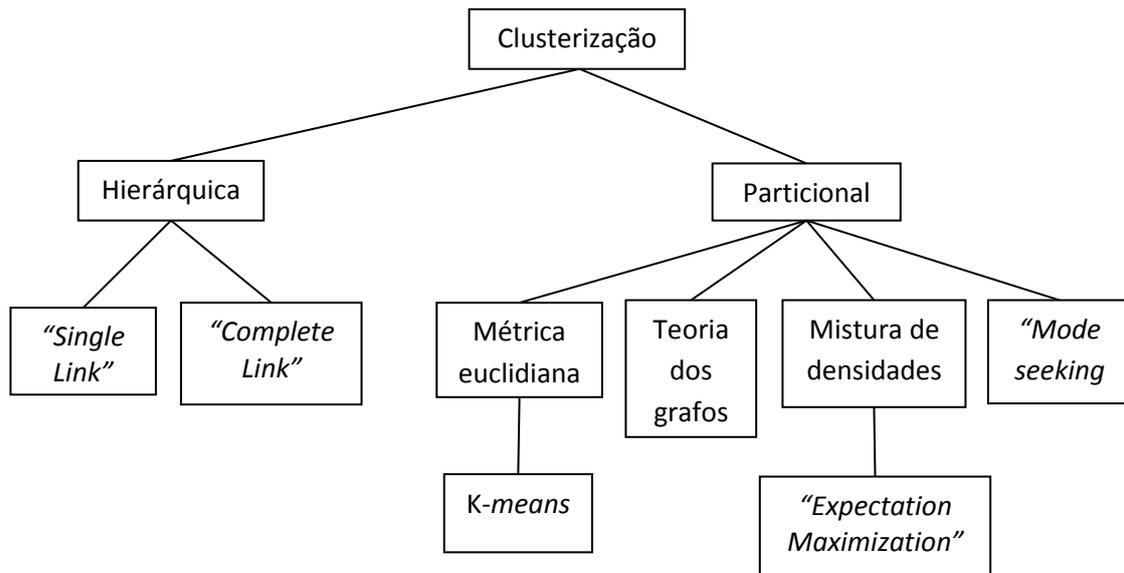


Figura 6. Exemplo de taxonomia para técnicas de *clusterização* [33].

O algoritmo *k-means* [29], classificado como hierárquico, é simples e um dos mais conhecidos na área de *clusterização*. Seu objetivo é particionar os dados em k *clusters*, tal que o número k seja previamente estabelecido. O algoritmo mais simples para o *k-means*, cujas entradas são um conjunto de vetores multidimensionais e o número k de *clusters*, está descrito abaixo:

1. Escolher aleatoriamente dentre os dados os k vetores que farão o papel de centroides, ou seja, serão os centros de cada um dos k *clusters*.
2. Calcular a distância entre cada vetor restante e os centroides. A distância euclidiana (descrita na seção Métricas e Medidas) é comumente utilizada nesta etapa.
3. Alocar cada vetor a um centroide usando como critério a menor distância entre eles, ou seja, o centroide que possuir menor distância em relação ao vetor irá alocá-lo para seu *cluster*. Geralmente a medida de distância utilizada é a distância euclidiana.
4. Recalcular os centroides de cada *cluster* baseado na média aritmética dos membros do *cluster*.
5. Repetir o passo 2.

6. Caso nenhum vetor tenha alterado seu *cluster* o algoritmo termina. Caso contrário, o algoritmo continua no passo 3.

O algoritmo *k-means* é normalmente relacionado aos mapas auto-organizáveis de Kohonen (SOM) devido seus objetivos semelhantes. O SOM é uma ANN especial para área de análise de dados porque pode ser utilizado para reduzir a dimensionalidade dos dados através de *clusterização* ou pode ser utilizado como uma boa alternativa de visualização gráfica das entradas. O SOM será detalhado na próxima seção.

2.4.3 Self-Organizing Maps (SOM)

A estrutura de um mapa auto-organizável [5], ilustrada na Figura 7, é composta por uma camada de entrada com vetores que representam os dados e uma camada de saída formada por neurônios disposto em uma grade que é normalmente uni ou bidimensional. Não há camadas intermediárias, como em outras redes neurais.

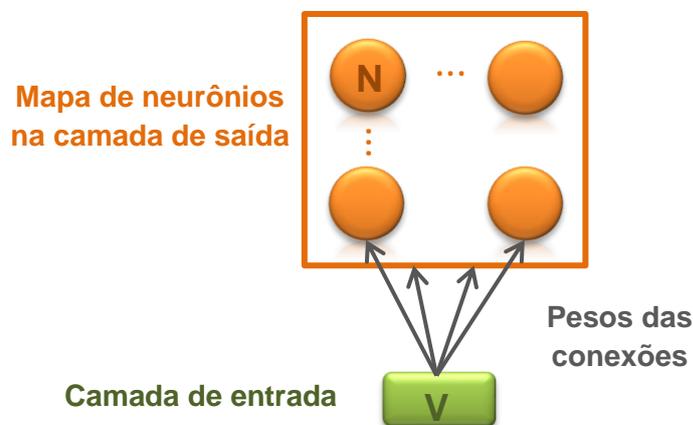


Figura 7. Estrutura de um mapa auto-organizável de Kohonen.

Cada neurônio está conectado com os vetores da camada de entrada e possui pesos que representam as sinapses dos neurônios biológicos [37]. A aprendizagem do SOM ocorre de forma não supervisionada e competitiva, ou seja, os neurônios competem entre si, disputando a oportunidade de responder às características dos dados de entrada, de maneira que apenas um estará ativo para um determinado vetor de entrada.

A Figura 8 descreve matematicamente a representação dos vetores de dados e do vetor de pesos dos neurônios. Um vetor de dados V é representado pelos atributos que lhe cabem, enquanto um neurônio é representado pelo vetor de pesos originados das conexões com cada neurônio. Cada peso w possui um índice referente à conexão do neurônio j com cada um dos vetores de dados.



$$V = [x_1, x_2, \dots, x_i]$$

$$N = [w_{j1}, w_{j2}, \dots, w_{ji}]$$

Figura 8. Representação dos vetores de dados e dos neurônios no SOM.

O processo de aprendizado de um SOM é baseado na topologia do córtex cerebral de animais desenvolvidos, em que entradas sensoriais diferentes são mapeadas para áreas diferentes e entradas semelhantes são agrupadas em áreas próximas. Para entender o funcionamento do SOM, o algoritmo de treinamento [5] [32][38][39] será descrito abaixo por etapas, considerando que uma unidade de tempo t corresponde também a uma iteração do algoritmo:

1. Inicializar com valores aleatórios o vetor de pesos para cada neurônio da grade. Ao longo do treinamento estes pesos serão ajustados para se aproximar dos dados de entrada.
2. Selecionar aleatoriamente um vetor de entrada v dentre o espaço de dados.
3. Calcular a diferença entre o vetor de dados v selecionado e cada um dos neurônios do mapa a fim de encontrar a menor distância. Este cálculo pode ser feito a partir de diversas métricas de distância entre vetores; geralmente é feito baseado na distância euclidiana (descrita na seção Métricas e Medidas).
4. Encontrar o neurônio vencedor ou BMU (do inglês *Best Matching Unit*). O neurônio vencedor é aquele que apresenta a menor distância para o vetor de dados apresentado.
5. Definir uma vizinhança topológica de neurônios tendo como centro o BMU. A vizinhança indica o nível de cooperação entre o BMU e seu vizinho j . A equação exponencial 4 representa a função de cooperação para neurônios

vizinhos. O parâmetro $d_{j,c}$, o caso de um mapa com duas dimensões, corresponde a distância euclidiana entre o BMU, representado pela letra c , e um neurônio vizinho j .

$$h_{j,c}(t) = \exp\left(-\frac{d_{j,c}^2}{2\sigma^2(t)}\right) \quad (4)$$

A equação 4 decresce no tempo assim como o raio da vizinhança σ , indicado pela equação 5. O parâmetro σ_0 é o raio inicial e τ é uma constante de tempo.

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\tau}\right) \quad (5)$$

Desta forma, quanto o tempo t aumenta, a largura do raio da vizinhança decresce a uma taxa exponencial e a função de cooperação diminui de maneira semelhante. A Figura 9 é um exemplo da vizinhança topológica diminuindo ao longo do tempo.

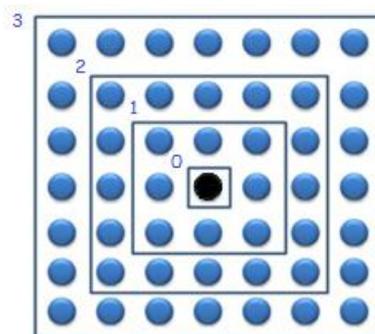


Figura 9. Vizinhança topológica de um neurônio [39].

6. Atualizar os pesos das conexões do neurônio BMU e de sua vizinhança segundo a equação 6.

$$w_j(t+1) = w_j(t) + \alpha(t) \cdot h_{j,c}(t) \cdot (v - w_j(t)) \quad (6)$$

O parâmetro $\alpha(t)$ é a taxa de aprendizagem da rede, que deve decrescer no tempo (mantendo-se entre 1 e 0), semelhante a função do raio da vizinhança. Desta equação entende-se que o peso de uma conexão na próxima iteração

do algoritmo sofrerá uma fração de perda ou esquecimento (α) em relação ao vetor de dados selecionado e o peso da iteração corrente.

7. Voltar ao passo 2 para apresentar as demais entradas.

Quando o SOM é aplicado a um grupo de dados o mapa topológico formado pode ser utilizado como ferramenta para visualizar os grupos de dados que foram aglomerados em áreas próximas, ou seja, é possível identificar *clusters* [28][29] no mapa.

2.4.4 K-means versus SOM

Ambas as técnicas *K-means* [29] e SOM [5] podem ser utilizadas para formar *clusters* com os dados de entrada e ambas empregam medidas de distância para alcançar seus objetivos, normalmente a distância euclidiana. Porém o número de *clusters* do *k-means* deve ser escolhido de acordo com o número de *clusters* que se deseja formar, enquanto no SOM o número de vetores de dados é independente do número de *clusters*.

O *k-means* é um algoritmo simples baseado apenas nas ideias de distância e média aritmética, já o SOM é uma técnica inteligente baseada no comportamento do córtex cerebral de animais desenvolvidos e que possui parâmetros mais complexos, tal qual a ideia de vizinhança. Embora exista essa diferença, ao executar o SOM com a vizinhança igual a zero, ele se torna um algoritmo convencional, semelhante ao *k-means*. Para a proposição da métrica deste trabalho, o SOM foi a tecnologia escolhida, já que é possível formar o mapa sem determinar uma quantidade fixa de *clusters*.

Capítulo 3

Proposição da Métrica

Este capítulo apresenta a formulação da métrica proposta para a avaliação do valor semântico entre/em sistemas complexos e descreve o relacionamento entre Semiótica, mapas auto-organizáveis de Kohonen e comunicação, conceitos presentes no capítulo anterior, que são fundamentais para a construção do modelo de métrica. O capítulo também explica a representação vetorial de conceitos objetivos e subjetivos, a comparação dos vetores através da rede de Kohonen e as interpretações dos resultados.

3.1 Visão Geral do Modelo

Cada ser humano, ao experimentar um processo de semiose, constrói sua representação do objeto através da significação inerente ao seu universo mental, de forma que a semiose em diferentes seres ocorre de maneira distinta e gera signos distintos.

Em um processo de comunicação um-a-um (um emissor e um receptor), quando o destinatário é um ser humano, e não necessariamente o emissor também, o sinal enviado não se limita a funcionar como um simples estímulo, mas solicita uma resposta interpretativa do destinatário [13]. Desta forma, ao se tratar de dois seres humanos, cada ser envolvido possui em sua mente uma interpretação, ou seja, um valor semântico para o objeto enviado ou recebido. Isto quer dizer que o significado de um objeto transmitido não se encontra no próprio objeto, mas na relação entre o emissor, o objeto, o receptor e o contexto [40].

Devido a esses fatores, Santaella diz que a semiótica não é uma chave milagrosa [9]. Já que sem conhecer o contexto sociocultural em que um signo se situa não se pode detectar as marcas que o contexto deixa na mensagem. Se o repertório de informações do receptor é muito baixo, a semiótica não o faz produzir interpretantes que vão além do senso comum.

Do ponto de vista prático, a completude do processo de comunicação ocorre quando o objeto é recebido. Porém, a qualidade do processo não depende exclusivamente do término do processo; na verdade a qualidade está intrinsecamente relacionada com a forma que o emissor codifica aquilo que deseja enviar e com a compreensão do receptor sobre o que está sendo percebido.

Tomando como exemplo um modelo semiótico de comunicação, ou seja, um modelo de comunicação que enfatiza a criação dos significados e a formação das mensagens a partir de signos [40], pode-se analisar as características mencionadas anteriormente quanto à fidelidade do processo comunicativo.

A Figura 10 ilustra um processo de comunicação simples com elementos semióticos. Este é um modelo semiótico de comunicação porque neste caso a comunicação não é considerada apenas um fluxo ou um processo, mas um sistema estruturado gerador de significado [20]. Note-se a pequena-grande divergência semântica entre os signos, que funcionalmente guardam similaridade (bola), entretanto são objetos diferentes no mundo real (bola de futebol e bola de *baseball*) e possuem diferença semântica.

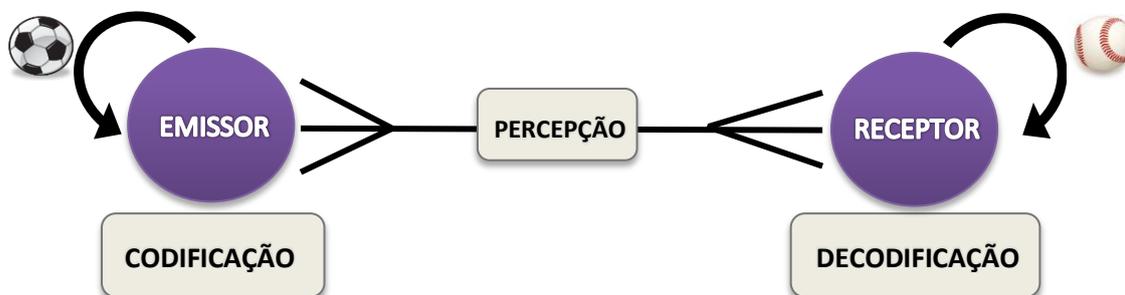


Figura 10. Exemplo de comunicação com semiose.

Lê-se a Figura 10 da seguinte maneira: o emissor codifica o objeto da comunicação da forma que acha conveniente, o receptor percebe o objeto, decodifica-o e dá sentido ao objeto de acordo com o conhecimento que já detém. Para este caso em específico, uma bola de futebol foi codificada com a palavra “bola”; ao perceber a palavra o receptor realiza uma associação com algo que já foi percebido anteriormente, que pode igualmente ser entendido pela palavra recebida. A intenção do emissor foi transmitir a informação de uma bola de futebol, porém o receptor talvez não conheça uma bola de futebol, mas conheça, por meio de alguma

percepção em um tempo anterior, uma bola de *baseball*. Assim, a bola do início do processo de comunicação difere em algumas características da bola do final. A comunicação não deixou de acontecer, porém não foi a mais fiel possível, já que o valor semântico da bola para o receptor em comparação com o emissor possui um grau de dissimilaridade.

Para que seja possível comparar os valores semânticos dados a um objeto por diferentes seres, o modelo de métrica proposta neste documento utiliza-se dos mapas auto-organizáveis de Kohonen, já que esta técnica de aprendizado não supervisionado permite organizar características semelhantes em um espaço 2D, alocando entradas próximas em locais próximos. Os mapas de Kohonen também permitem calcular a distância espacial entre as entradas apresentadas à rede e conseqüentemente à distância conceitual entre as mesmas, ou seja, uma medida de similaridade ou dissimilaridade, já que quanto menor é a distância espacial menor será a distância conceitual.

O modelo de métrica de avaliação do valor semântico entre/em sistemas proposto neste capítulo leva em consideração os elementos que envolvem o processo semiótico que ocorre nas mentes de seres humanos, compreendendo a desconstrução de objetos em suas características mais objetivas como também naquelas mais subjetivas, representando-as de maneira semelhante para que um objeto seja a junção de seus segmentos objetivo e subjetivo.

O modelo de métrica pode ser visualizado na Figura 11. As etapas necessárias ao processo são:

- i) Codificar os vetores de atributos que representarão um conceito;
- ii) Treinar a rede de Kohonen com os vetores de atributos gerados por representantes do grupo emissor de um conceito;
- iii) Testar os vetores de atributos gerados por representantes do grupo receptor de um conceito;
- iv) Calcular a distância entre os vetores que se deseja comparar, interpretando a distância espacial calculada como uma distância também semântica.

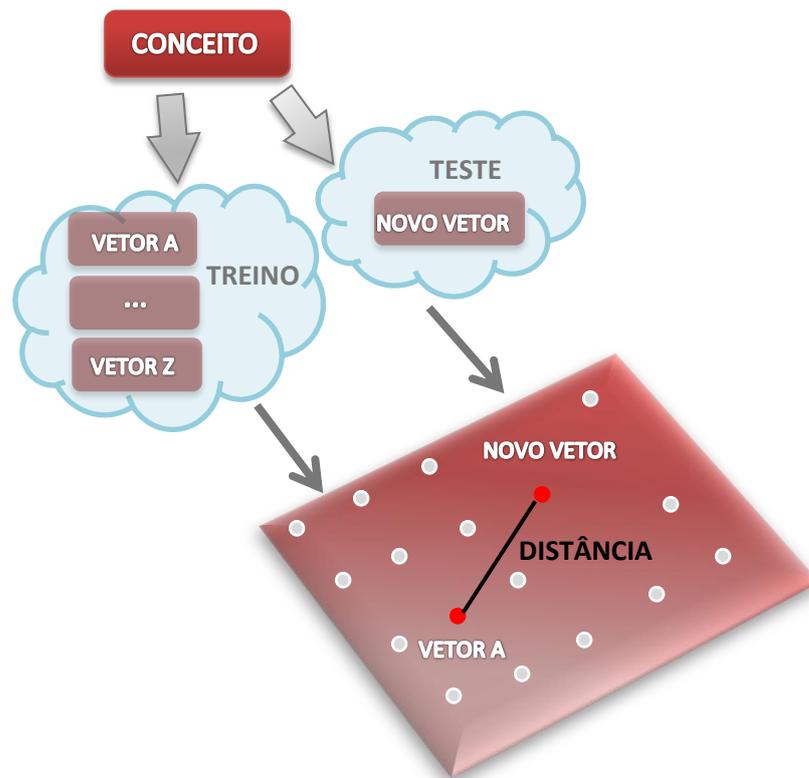


Figura 11. Modelo da métrica proposta.

3.2 Representação Vetorial do Conceito

Representar conceitos ou conhecimentos que habitam a mente de um ser humano é uma tarefa muito complexa. Independente da divergência científica para definir a mente [41], é unânime a dificuldade de penetrar no espaço mental e descobrir como uma determinada pessoa entende conceitos, o que eles passam a significar e as lembranças e sentimentos que provocam.

Alguns padrões humanos facilitam a definição de conceitos objetivos, por exemplo, é possível contar o número de palavras em um texto, definir uma imagem através de um sistema de cores [42] como o RGB (do inglês, *Red, Green and Blue*), medir e pesar através do sistema internacional de unidades (SI – do francês *Système International D'Unités*) [43], assim como definir formas geométricas básicas como quadrado, círculo e cubo. Porém, ao tentar representar conceitos subjetivos, tais como beleza, dificuldade e qualidade esbarra-se na falta de padronização, já que para diferentes pessoas os conceitos variam. Ainda assim, ao comparar a

representação do subjetivo para duas pessoas é possível utilizar escalas de medidas [29] para suas respostas, limitando o senso de julgamento de maneira igualitária. Se a pergunta “Você está feliz?” fosse feita para duas pessoas seria possível limitar as respostas em uma escala com quatro alternativas, tais como “Sim, muito feliz”, “Sim, estou feliz”, “Sinto-me normal” e “Não estou feliz”.

Para a métrica proposta os vetores de entrada para o mapa de Kohonen podem ser formados apenas por elementos subjetivos ou por uma composição de elementos objetivos e subjetivos, como mostra a Figura 12.

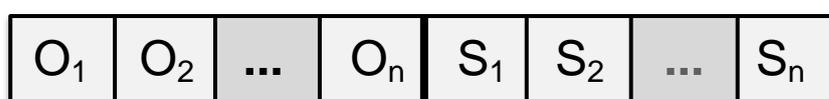


Figura 12. Vetor de conceitos objetivos e subjetivos.

Tomando como exemplo uma camisa comum, é possível desconstruí-la em um vetor de atributos objetivos e subjetivos, exemplificados na segunda coluna da Tabela 2.

A terceira coluna exemplifica uma forma de padronizar as respostas para atributos objetivos e limitar as respostas para atributos subjetivos.

Tabela 2. Atributos de uma camisa.

OBJETIVO	Cor	(R,G,B)
	Manga	Sim ou Não
	Tamanho	P, M, G, GG
SUBJETIVO	Qualidade do tecido	Muito bom, bom, regular, ruim
	Beleza	Muito bonita, bonita, comum, feia
	Custo/ Benefício	Muito bom, bom, regular, ruim

Ao construir o vetor da camisa representada pela Figura 13, obtemos o vetor da Tabela 3. Os três atributos classificados como objetivos são simples de definir, pois são padronizados. Desta forma duas pessoas provavelmente definiriam estes atributos de maneira igual, porém os três últimos dependem do que qualidade, beleza e custo/benefício significa para cada pessoa.

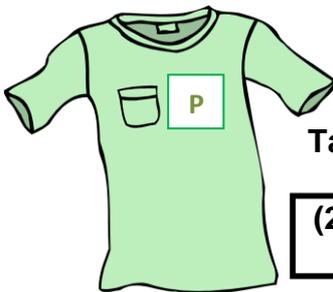


Tabela 3. Vetor representativo de uma camisa.

(202, 255, 202)	Sim	P	?	?	?
--------------------	-----	---	---	---	---

Figura 13. Camisa.

A escolha dos atributos que compõem o vetor de conceitos de um objeto é essencial para o processamento e a interpretação dos resultados gerados após a utilização dos dados no modelo proposto para a métrica. A camisa do exemplo foi desconstruída em apenas seis atributos, porém, à medida que este número aumenta mais detalhes sobre a camisa ficam evidentes. A complexidade do processamento no mapa de Kohonen também cresce, visto que mais atributos devem ser comparados para encontrar o neurônio vencedor quando uma entrada é apresentada à rede.

3.3 Comparação e Interpretação

Para comparar a perspectiva de um objeto para pessoas diferentes é necessário que para cada pessoa seja associado um vetor de atributos do objeto em questão. Os dados gerados de acordo com as pessoas classificadas como emissoras de uma mensagem são usados para o treinamento do mapa de Kohonen, os dados dos receptores da mensagem são apresentados à rede após o treinamento, na fase de testes. Desta forma é possível saber em que região do mapa a nova entrada se encaixa melhor, visto que o mapa de Kohonen compara

uma nova entrada com os neurônios já presentes no mapa através de uma medida de distância.

A distância entre dois elementos no mapa de Kohonen reflete o grau de similaridade entre dois vetores. Portanto, para medir quão próximo está o valor semântico de um objeto para uma pessoa A do valor dado pela pessoa B através do modelo proposto neste trabalho basta encontrar a distância entre os dois vetores conceituais no mapa 2D.

Argumentamos que se o sistema emissor de uma mensagem obtiver o conhecimento do grau de entendimento do receptor, o processo de interpretação da mensagem poderia ser adaptado objetivando a completude do sentido da mensagem. Em termos semióticos, aproximar a semântica de um signo ou conjunto de signos para um intérprete X em relação a um intérprete Y diminuiria a disparidade de entendimento ao final do processo de comunicação.

Capítulo 4

Experimentos e Resultados

Este capítulo contém os experimentos realizados para validar o modelo de métrica proposto neste trabalho. Dois domínios de conhecimentos foram escolhidos para a validação: o suporte à decisão e educação.

4.1 Cenário Um: Suporte à Decisão

O cenário de suporte a decisão utilizou dados da dissertação de mestrado de Bernardo Caldas [44] como base, pois os dados recolhidos neste trabalho além de acessíveis se adequam a proposta do teste da métrica em suporte à decisão.

Bernardo Caldas, em sua dissertação, desenvolveu um sistema de apoio a decisões estratégicas em empresas incorporando conhecimentos de inteligência artificial. Para validar o algoritmo criado no sistema, Bernardo construiu uma metodologia baseada em Sistemas Imunológicos Artificiais para aplicar formulários a executivos decisores e funcionários subordinados de 13 empresas comerciais privadas de produtos de informática da cidade do Recife, Pernambuco, Brasil durante os anos de 2007 e 2008.

Os formulários recolhidos refletiram a opinião de cada pessoa sobre o questionamento “Quais informações você julga apropriadas para apoiar a tomada de decisão da compra de uma nova plataforma tecnológica para sua organização?”. Tanto decisores quanto subordinados responderam a pergunta de acordo com vinte atributos formulados durante o processo de coleta de informações como fruto do acordo entre o pesquisador e os decisores. Os vinte conceitos são detalhados na Tabela 4.

Cada empresa foi representada por três indivíduos, um decisor e dois subordinados. Cada um dos envolvidos respondeu o questionamento optando no máximo por 10 dos 20 atributos existentes e definindo a ordem de prioridade entre

as 10 escolhas. As respostas de decisores e subordinados de cada empresa são apresentadas em sua totalidade no Anexo A.

Tabela 4. Atributos subjetivos de uma decisão de compra

	Atributo	Significado do atributo
1	Benefícios	Benefícios auferidos com a aquisição e implantação do projeto.
2	Consultoria	Contratação de empresa especializada para apoiar a decisão.
3	Custo	Custo final do projeto (da solução).
4	Entrega	Cumprimento do prazo de entrega dos produtos ou serviços.
5	Especificações	Atendimento de todas as especificações do projeto.
6	Fluxo	Existência dos recursos financeiros internos para honrar os pagamentos.
7	Fornecedor	Idoneidade e tradição na fabricação dos produtos ou prestação de serviços.
8	Garantia	Garantia dos produtos e serviços comercializados pelo fornecedor.
9	Mercado	Receptividade do mercado quanto à tecnologia para aquisição.
10	Necessidade	Percepção do empresário da real necessidade de mudança em sua organização.
11	Padrões	Os produtos obedecem a padrões internacionais de fabricação.
12	Pagamento	Os pagamentos estão alinhados ao fluxo financeiro da organização.
13	Preço	O preço cobrado é justo e compatível com outros fornecedores.
14	Qualidade	O fornecedor mantém certificação ISO (do inglês <i>International Organization for Standardization</i>) de Gestão da Qualidade na produção de produtos ou serviços prestados.
15	Relacionamento	Nível de relacionamento entre a organização e o fornecedor.
16	Retorno	Retorno financeiro do investimento realizado.
17	Serviços	Pré-venda. O fornecedor presta serviços de apoio no levantamento e especificação de produtos e serviços aplicáveis.

4.1.1 Representação da Decisão de Compra

Para realizar o experimento com a métrica proposta. 10 das 13 empresas foram utilizadas como dados de entrada para os mapas auto-organizáveis de Kohonen. Para tal, os atributos nominais foram convertidos em numéricos e a ordem

de prioridade estabelecida por cada indivíduo será levada em consideração também de forma numérica.

Os 20 atributos detalhados anteriormente são representados pelos números naturais de 1 a 20, ou seja, os números apresentados na Tabela 4. Ao preparar os dados que serão usados no mapa de Kohonen é necessário não só transformá-los para uma forma numérica como também definir uma escala padrão para os valores, de forma que nenhum dos valores se sobressaia a outro ou os resultados sofram alguma tendência. Assim, os atributos foram normalizados em uma escala entre zero e um, como indica a equação 7.

$$\text{Novo atributo} = \text{Atributo}/20 \quad (7)$$

Desta forma, os atributos serão representados pelos novos valores normalizados entre zero e um, conforme mostra a Tabela 5.

Tabela 5. Normalização dos atributos

Atributo	1	2	3	4	5	6	7	8	9	10
Novo atributo	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
Atributo	11	12	13	14	15	16	17	18	19	20
Novo atributo	0,55	0,6	0,65	0,7	0,75	0,8	0,85	0,9	0,95	1

Já definidos os valores que serão utilizados para cada um dos vinte atributos, resta definir os valores que representarão as prioridades referentes a cada um dos atributos escolhidos por cada pessoa que respondeu o formulário. Já que foi fixado no formulário que cada pessoa poderia optar por no máximo 10 atributos, também serão 10 as prioridades.

Analogamente aos atributos, as dez prioridades de 1 a 10 serão representadas por um peso normalizado entre 0 e 1, cujo maior peso corresponde a maior prioridade. Logo, sendo 10 prioridades, o maior peso será representado pelo valor 1 e o menos pelo valor 0,1, conforme mostra a Tabela 6.

Tabela 6. Valores normalizados para as prioridades

Prioridade	1º	2º	3º	4º	5º	6º	7º	8º	9º	10º
Nova prioridade	1,0	0,9	0,8	0,7	0,6	0,5	0,4	0,3	0,2	0,1

Para demonstrar as normalizações de atributos e prioridades, a primeira empresa presente no anexo será tomada como exemplo, como mostra a Tabela 7.

Tabela 7. Resposta do decisor (D) e dois subordinados (S1 e S2)

	1º	2º	3º	4º	5º
D1	Especificações	Serviços	Qualidade	Preço	Pagamento
S1	Serviços	Especificações	Preço	Qualidade	Pagamento
S2	Necessidade	Consultoria	Retorno	Fornecedor	Serviços
	6º	7º	8º	9º	10º
D1					
S1					
S2	Especificações	Preço	Qualidade	Pagamento	Suporte

Aplicando apenas ao decisor da Tabela 7 os passos de representação e normalização explicados anteriormente obtemos a Tabela 8.

Tabela 8. Representação do vetor do decisor (D1)

VETOR DO DECISOR										
Atributo	0,25	0,85	0,7	0,65	0,6	0	0	0	0	0
Prioridade	1,0	0,9	0,8	0,7	0,6	0	0	0	0	0

De forma semelhante, as opiniões dos demais indivíduos serão associadas a um vetor composto pelos atributos escolhidos e suas respectivas prioridades. A totalidade dos dados está exposta no Apêndice A, onde constam 10 empresas, totalizando 30 opiniões (10 de decisores e 20 de subordinados).

4.1.2 Treinamento e Teste

Existem diversas ferramentas que proporcionam o uso do algoritmo do mapa auto-organizável de Kohonen, tal como o *software* Matlab [45], através de um *ToolBox* próprio para SOM [46]. Também existem bibliotecas em várias linguagens de programação que podem ser utilizadas, por exemplo, a JML (*Java Machine*

Learning Library) [47]. Cada ferramenta ou biblioteca possui vantagens e desvantagens em relação às características do SOM. Tanto o Matlab como a JML foram testadas, porém uma terceira alternativa mostrou-se satisfatória do ponto de vista gráfico para o propósito deste trabalho, o *software* Spice-SOM [48], criado por Cao Thang, visto que essa ferramenta permite mostrar *labels* que representam as entradas fornecidas à rede.

O *software* Spice-SOM [48] é simples, porém proporciona a alteração de vários parâmetros associados ao SOM e mostra graficamente o erro ao longo do tempo. Para utilizar a ferramenta é necessário importar os dados de entrada de acordo com o modelo aceito pela ferramenta, depois basta configurar os parâmetros do mapa auto-organizável de Kohonen para que a rede seja treinada e o mapa seja gerado.

Os vetores de dados de entrada (*i.e.* opiniões dos subordinados e opiniões dos decisores) receberam um *label* cada um em que ‘E’ significa a empresa, ‘S’ o subordinado e ‘D’ o decisor.

Após diversos testes com a ferramenta e a observação do erro ao longo do tempo a configuração do mapa de Kohonen com a dimensão 7x7 mostrou-se satisfatório em relação à quantidade de vetores de entrada, visto que com o valor 7 as entradas não se espalham tanto ao longo do mapa deixando muitos neurônios vazios e também não se aproximam muito de forma a prejudicar a identificação de regiões com *clusters* formados.

A Figura 14 mostra o mapa de Kohonen gerado pela ferramenta Spice-SOM. Os pontos com *label* na forma “EnSm” representam as opiniões dos subordinados e os pontos com *label* na forma “EnDm” representam as opiniões dos decisores. Os três pontos iniciados por “ER” são vetores criados como referência, ou seja, não estavam na base de dados original; os vetores de opiniões dos subordinados e do decisor foram criados propositalmente com os valores idênticos para que fossem observadas suas posições no mapa, que deve ser a mesma.

O mapa gerado mostrou que é possível comparar o valor semântico atribuído a uma decisão de compra por vários indivíduos, que muitas vezes possuem o mesmo ambiente de convivência, simplesmente criando os vetores de atributos

subjetivos de forma cuidadosa para agregar o maior número de informações relevantes. A partir do mapa resultante pode-se calcular a distância entre quaisquer neurônios representativos das entradas.



Figura 14. SOM gerado pela ferramenta Spice-SOM

Tomando como exemplo os dois vetores selecionados em vermelho. O neurônio de cima, cuja coordenada espacial é igual a (5,1), representa a opinião do segundo funcionário subordinado da sétima empresa. Já o segundo neurônio, cuja coordenada espacial é igual a (6,2), corresponde as opiniões do primeiro subordinado da oitava empresa e ao decisor da primeira empresa. Sabendo o valor das coordenadas é possível calcular a distância euclidiana entre os neurônios:

$$D \text{ euclidiana} = \sqrt{(5 - 6)^2 + (1 - 2)^2}$$

$$D \text{ euclidiana} = \sqrt{2}$$

O valor $\sqrt{2}$ representa não só a distância espacial entre os neurônios como também a distância semântica do conceito de decisão de compra para o subordinado da oitava empresa em relação aos outros dois indivíduos.

4.2 Cenário Dois: Educação

A sala de aula é um bom exemplo de interação e comunicação entre sistemas complexos. O conhecimento que o professor detém é lançado aos alunos através de diversos meios e utilizando signos ou conjunto de signos tais como palavras, imagens, gestos, sons, cores, etc. Mudando o ambiente de aprendizagem da sala de aula para o ensino a distância (EaD) [49] ou o ensino através de sistemas tutores inteligentes (ITSs, do inglês *Intelligent Tutoring Systems*) [50] ou ainda cursos *on-line*, a comunicação tende a ficar mais complexa, visto que na sala de aula é o professor quem compreende os alunos e sabe reconhecer as melhores maneiras de se aplicar este ou aquele recurso [51]; já nas demais abordagens de ensino a presença do professor é limitada ou eliminada.

O entendimento dos signos por parte dos alunos depende dos mais variados fatores, pois um aluno aprende dentro dos limites do seu conhecimento e estado emocional. Além de que o próprio signo pode não ter sido a melhor escolha para representar o que deveria ser transmitido. O processo de aquisição de conhecimento envolve diferentes sistemas complexos em comunicação e a característica fundamental de um processo de comunicação bem executado é a igualdade entre a mensagem (signo) transmitida e a recebida.

Softwares que se propõem a orientar a aquisição de conhecimento podem utilizar a métrica como um componente a mais de sua arquitetura a fim de individualizar o envolvimento do aluno. Aproximar o que deve ser transmitido e o que realmente foi entendido pelo aluno é fundamental para que o ciclo educacional seja satisfatório. Isto significa utilizar a métrica de avaliação semântica como medida para o *feedback* do processo cíclico da aprendizagem.

4.2.1 Cenário e Base de Dados

Não foi possível para este trabalho desenvolver um teste completo da métrica com um cenário de dados sobre Educação, visto que não foi encontrada uma base de dados já existente adequada.

Entretanto, argumentamos que a análise da semiótica e da comunicação no domínio de educação é fundamental para realizar avaliações de professores, alunos, métodos de aprendizagem e *softwares* de ensino. Nesse sentido, quando um tutor, seja ele um ser humano ou um *software*, transmite conhecimento para um aluno sempre há um objetivo associado ao processo de comunicação; além de fazer com que o aluno receba e entenda a informação passada, há também a preocupação que o resultado esperado em níveis de aprendizagem seja despertado na mente do estudante, fazendo com que habilidades e competências sejam desenvolvidas a partir do novo conhecimento adquirido. Por analogia ao exemplo anterior, defendemos que é possível a utilização da métrica proposta, desde que os conceitos a serem medidos tenham a definição de atributos objetivos e subjetivos realizada apropriada e oportunamente.

Para que seja possível avaliar um aluno normalmente são aplicadas provas ou cobrados projetos. Utilizar a métrica proposta para avaliar as evidências de aprendizagem de um aluno em comparação com a expectativa do tutor é uma alternativa não só possível, bem como interessante de avaliação. Assim, seria possível analisar o nível de entendimento de um estudante através de conceitos subjetivos e realizar *feedbacks* de forma mais próxima da realidade e até acompanhar a evolução do conhecimento ao longo do tempo. Também seria possível recuperar informações subjetivas de alunos para tomar decisões e agir. Por exemplo, nivelar o nível de dificuldade de avaliações objetivas de acordo com o conceito subjetivo de dificuldade para os próprios alunos.

Capítulo 5

Conclusões e Trabalhos Futuros

Este último capítulo destina-se às considerações finais do trabalho, tanto sobre o que foi realizado como o que pode ser melhorado e incorporado futuramente.

5.1 Conclusões

Este trabalho teve o objetivo principal de desenvolver uma métrica para avaliação do valor semântico dado a um signo por um sistema e a diferença ao comparar com o valor dado por outro sistema. Empregando semiótica e computação foi possível idealizar um modelo para a métrica, viabilizando, por meio de distâncias matemáticas simples, a similaridade ou dissimilaridade entre o entendimento de um conceito por diferentes sistemas.

Os dois cenários propostos (um experimental – decisões comerciais e outro argumentativo – educação) para realizar os experimentos com a métrica refletem ambientes complexos de interação e comunicação entre seres humanos, o que torna complicada a tarefa de tentar objetivar conceitos subjetivos. Porém, através de dados controlados e limitados a um escopo foi viável por em prática a subjetividade como entrada para uma rede neural no primeiro cenário de testes. Não foi possível montar uma base de dados apropriada para o segundo cenário, porém este é um ponto importante a ser explorado no futuro.

A métrica foi desenvolvida para ser um componente de sistemas maiores, inseridos em diversos domínios, tais como os dois cenários apresentados e contribui para a avaliação da semântica associada a um signo por um sistema e para a comparação de conceitos subjetivos entre sistemas em comunicação. A métrica pode ser aplicada como informação de maior valor agregado para tomada de decisões e atuação no ambiente a que é empregada, de acordo com o domínio em questão.

5.2 Trabalhos Futuros

Durante o desenvolvimento do trabalho algumas ideias surgiram para incrementar o seu desempenho e aplicabilidade no futuro. As ideias estão descritas abaixo:

- Desenvolver uma aplicação própria para os mapas auto-organizáveis de Kohonen e perder a dependência com as restrições das ferramentas prontas. Desta forma a métrica poderá ser testada de maneira mais controlada, utilizando o algoritmo de Kohonen e a visualização dos mapas de forma mais completa.
- Aprimorar a identificação de *clusters*, utilizando algoritmos de clusterização, como o *k-means*, após a construção do mapa de Kohonen. Desta forma, a formação de *clusters* não será apenas visual e sim automática de acordo com um algoritmo.
- Investigar a incorporação da lógica fuzzy como forma de aproximar a métrica proposta da incerteza humana diante de conceitos. Já que a lógica fuzzy pode auxiliar na representação de conceitos não quantificáveis.
- Investigar o uso de outras medidas de distância ao invés da distância euclidiana e as consequências de uma troca.
- Buscar ou construir bases de dados mais complexas e inseridas em contextos humanos comuns, como dentro do domínio de educação já citado.

Bibliografia

- [1] PEIRCE, C. S. **Semiótica**. Tradução de José T. C. Neto. 4. ed. São Paulo: Perspectiva, v. Coleção Estudos, 2012.
- [2] ECO, U. **Signo**. Tradução de Francisco Serra Cantarell. 2ª. ed. [S.l.]: [s.n.], 1994.
- [3] MORRIS, C. **Fundamentos da Teoria dos Signos**. Tradução de António Fidalgo. [S.l.]: [s.n.].
- [4] SANTAELLA, L. **O Que é Semiótica**. São Paulo: Brasiliense, 1983.
- [5] KOHONEN, T. The Self-Organizing Map. **Proceedings of the IEEE**, 78, Setembro 1990. 1464-1480.
- [6] CHARLES Sanders Peirce. **Wikipedia**. Disponível em: <http://en.wikipedia.org/wiki/Charles_Sanders_Peirce>. Acesso em: Outubro 2013.
- [7] FERDINAND de Saussure. **Wikipedia**. Disponível em: <http://en.wikipedia.org/wiki/Ferdinand_de_Saussure>. Acesso em: Novembro 2013.
- [8] NÖTH, W. **Panorama da Semiótica: de Platão a Peirce**. São Paulo: Annablume, 1995.
- [9] SANTAELLA, L. **Semiótica Aplicada**. São Paulo: Cengage Learning, 2012.
- [10] LOCKE, J. Book IV: Knowledge. In: _____ **An essay concerning human understanding**. [S.l.]: [s.n.], Cap. xxi. Disponível em <http://www.earlymoderntexts.com/lo4.html>.
- [11] SAUSSURE, F. D. **Curso de Linguística Geral**. São Paulo: Cultrix, 2006.
- [12] CHANDLER, D. **Semiotics: The Basics**. 2ª. ed. [S.l.]: Routledge, 2007.
- [13] ECO, U. **Tratado de Semiótica Geral**. Tradução de Carlos Manzano. 5ª. ed. [S.l.]: Lumen, 2000.
- [14] GUDWIN, R. R. **Semiônica: Uma Proposta de Contribuição à Semiótica Computacional**. Universidade Estadual de Campinas (UNICAMP). [S.l.]. 2002.
- [15] BARROS, R. C. D. **Estudos Preliminares de Semeiose Automática no Subsistema de Percepção de Agentes Inteligentes**. Universidade de Pernambuco (UPE). [S.l.]. 2009.
- [16] QUEIROZ, J. **Semeiose Segundo C. S. Peirce**. São Paulo: EDUC, 2004.

- [17] GILL, D.; ADAMS, B. **ABC of Communication Studies**. 2^a. ed. [S.I.]: [s.n.], 1998.
- [18] SOUSA, J. P. **Elementos de Teoria e Pesquisa da Comunicação e dos Media**. 2^a. ed. [S.I.]: [s.n.], 2006.
- [19] DANCE, F. The "Concept" of Communication. **Journal of Communication**, 1970.
- [20] FISKE, J. **Introduction to Communication Studies**. 2^a. ed. [S.I.]: Routledge, 1990.
- [21] SHANNON, C.; WEAVER, W. **The Mathematical Theory of Communication**. [S.I.]: [s.n.]. University of Illinois Press.
- [22] JAKOBSON, R. Closing Statement: Linguistics and Poetics. *Style in Language*. ed. [S.I.]: Thomas Sebeok, 1960. p. 350-355.
- [23] BERLO, D. K. **O Processo da Comunicação: Introdução à Teoria e à Prática**. [S.I.]: Martins Fontes, 2003.
- [24] FLOYD, T. **Sistemas Digitais: Fundamentos e Aplicações**. 9^a. ed. [S.I.]: Bookman, 2007. Cap. 2.
- [25] KUROSE, J.; ROSS, K. **Redes de Computadores e a Internet: Uma Abordagem Top-Down**. 5^a. ed. São Paulo: [s.n.], 2010.
- [26] IEEE. **IEEE Standard Glossary of Software Engineering Terminology**. [S.I.]. 1990. Std 610.12.
- [27] DEZA, M. M.; DEZA, E. **Encyclopedia of Distances**. [S.I.]: Springer, 2009.
- [28] EVERITT, B. et al. **Cluster Analysis**. 5^a. ed. [S.I.]: Wiley, 2011.
- [29] GAN, G.; MA, C.; WU, J. **Data Clustering: Theory, Algorithms, and Applications**. [S.I.]. 2007.
- [30] INTERVIEW with Teuvo Kohonen. **IEEE Computational Intelligence Magazine**, v. 3, p. 4-5, 2008. CIM Editorial Officer.
- [31] VALENÇA, M. **Aplicando Redes Neurais: Um Guia Completo**. 1^a. ed. [S.I.]: Livro Rápido - Elógica, 2005.
- [32] HAYKIN, S. **Redes Neurais: Princípios e Prática**. Tradução de Paulo Martins Engel. 2^a. ed. Porto Alegre: Bookman, 2001.
- [33] JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data Clustering: A Review. **ACM Computing Surveys (CSUR)**, 31, 1999. 264-323.

- [34] JAIN, A. K.; DUBES, R. C. **Algorithms for Clustering Data**. [S.l.]: Prentice Hall, 1988. Michigan State University.
- [35] YUAN, B.; GEORGE, K. **Fuzzy Sets and Fuzzy Logic: Theory and Applications**. [S.l.]: Prentice Hall, 1995.
- [36] JONG, K. A. D. **Evolutionary Computation**. [S.l.]: [s.n.], 2006. Massachusetts Institute of Technology.
- [37] SYNAPSE. **Wikipedia**. Disponível em: <<http://en.wikipedia.org/wiki/Synapse>>. Acesso em: Novembro 2013.
- [38] BECKER, S. Unsupervised Learning Procedures for Neural Networks. **The international Journal of Neural Systems**, 1 e 2, 1991. 17-33.
- [39] DANTAS, A. B. D. S. **Sistema de Recomendação para Clientes de Vídeo Locadoras Baseado em Redes SOM**. Universidade de Pernambuco. [S.l.]. 2009.
- [40] FIDALGO, A. **Semiótica geral**. [S.l.]: [s.n.], 1999. Disponível em http://www.bocc.ubi.pt/_listas/tematica.php?codtema=12.
- [41] TEIXEIRA, J. F. **O Que é Filosofia da Mente**. [S.l.]: Brasiliense, v. Coleção Primeiros Passos, 1994. Disponível em <http://www.filosofiadamente.org/>.
- [42] GONZALEZ, R. C.; WOODS, R. E. Digital Image Processing. 2ª. ed. [S.l.]: Prentice Hall, 2002. Cap. 6.
- [43] THE International System of Units (SI). **Bureau International des Poids et Mesures (BIPM)**. Disponível em: <<http://www.bipm.org/en/si/>>. Acesso em: Novembro 2013.
- [44] CALDAS, B. J. D. B. **Suporte a Decisão Executiva Apropriada Baseado em Sistemas Imunológicos Artificiais (SSDeSIA)**. Universidade de Pernambuco. [S.l.]. 2008.
- [45] MATLAB. **MathWorks**. Disponível em: <<http://www.mathworks.com/products/matlab/>>.
- [46] SOM ToolBox. **Laboratory of Computer and Information Science (CIS)**. Disponível em: <<http://www.cis.hut.fi/somtoolbox/>>. Acesso em: Novembro 2013.
- [47] ABEEL, T.; DE PEER, Y. V.; SAEYS, Y. Java-ML: A Machine Learning Library. **Journal of Machine Learning Research**, 10, 2009. 931-934. Software disponível em: <http://java-ml.sourceforge.net/>.
- [48] THAN, C. Spice-SOM. Disponível em: <<http://download.cnet.com/windows/>>. Tutorial:

https://www.academia.edu/4207931/SELF_ORGANIZING_MAPS_SPICE-SOM_USERS_GUIDE.

- [49] ALVES, L. Educação a distância: conceitos e história no Brasil e no mundo. **Revista Brasileira de Aprendizagem Aberta e a Distância (RBAAD)**, v. 10, p. 83-92, 2011.
- [50] NWANA, H. S. Intelligent Tutoring Systems: an overview. **Artificial Intelligence Review**, 1990. 251-277.
- [51] Pesquisa sobre o Uso das Tecnologias de Informação e Comunicação nas Escolas Brasileiras. Comitê Gestor da Internet no Brasil (CGI). São Paulo. 2012.

Anexo A

Decisões apropriadas de decisores e subordinados

Este anexo apresenta os dados coletados na dissertação de Bernardo Caldas [44]. A partir dos formulários aplicados a decisores (D) e subordinados (S1 e S2) das 13 empresas, apenas as 10 primeiras empresas compõem a tabela abaixo.

EMPRESA1					
D	Especificações	Serviços	Qualidade	Preço	Pagamento
S1	Serviços	Especificações	Preço	Qualidade	Pagamento
S2	Necessidade	Consultoria	Retorno	Fornecedor	Serviços
	Especificações	Preço	Qualidade	Pagamento	Suporte
EMPRESA2					
D	Necessidade	Fluxo	Qualidade	Especificações	Preço
	Pagamento	Fornecedor			
S1	Custo	Benefícios	Necessidade	Entrega	Garantia
	Qualidade	Suporte	Tecnologia		
S2	Preço	Necessidade	Qualidade	Pagamento	Fornecedor
	Especificações	Fluxo			
EMPRESA3					
D	Qualidade	Especificações	Preço	Garantia	Suporte
	Entrega	Fornecedor	Pagamento	Benefícios	
S1	Benefícios	Preço	Especificações	Fornecedor	Qualidade
	Pagamento	Suporte	Entrega	Garantia	
S2	Necessidade	Consultoria	Fluxo	Custo	Benefícios
	Especificações	Suporte	Padrões	Retorno	Garantia
EMPRESA4					
D	Qualidade	Fornecedor	Preço	Suporte	Pagamento
	Garantia	Entrega			
S1	Benefícios	Fornecedor	Suporte	Custo	Qualidade
	Preço	Pagamento	Garantia		
S2	Fornecedor	Garantia	Suporte	Qualidade	Preço
	Pagamento	Entrega			
EMPRESA5					
D	Necessidade	Consultoria	Preço	Suporte	Tecnologia

S1	Benefícios	Retorno	Tecnologia	Custo	Preço
	Consultoria	Suporte			
S2	Necessidade	Tecnologia	Preço	Consultoria	Suporte
EMPRESA6					
D	Especificações	Necessidade	Entrega	Custo	Qualidade
	Relacionamento	Suporte	Benefícios	Preço	Pagamento
S1	Necessidade	Benefícios	Especificações	Suporte	Qualidade
	Custo	Pagamento	Preço	Entrega	Relacionamento
S2	Consultoria	Necessidade	Benefícios	Serviços	Especificações
	Qualidade	Garantia	Suporte	Custo	Retorno
EMPRESA7					
D	Necessidade	Custo	Fornecedor	Preço	Pagamento
	Retorno	Benefícios			
S1	Retorno	Custo	Fluxo	Fornecedor	Suporte
	Mercado	Necessidade	Entrega	Especificações	Garantia
S2	Necessidade	Benefícios	Retorno	Custo	Fornecedor
	Pagamento	Preço			
EMPRESA8					
D	Fornecedor	Mercado	Padrões	Qualidade	Especificações
	Necessidade	Entrega	Garantia	Preço	Pagamento
S1	Especificações	Garantia	Preço	Entrega	Pagamento
S2	Especificações	Garantia	Qualidade	Preço	Entrega
	Custo	Pagamento			
EMPRESA9					
D	Mercado	Tecnologia	Necessidade	Custo	Retorno
	Fornecedor	Relacionament o			
S1	Necessidade	Mercado	Custo	Retorno	Tecnologia
	Relacionamento	Fornecedor			
S2	Necessidade	Mercado	Benefícios	Custo	Retorno
	Tecnologia	Especificações	Entrega	Preço	Necessidade
EMPRESA10					
D	Especificações	Necessidade	Fornecedor	Qualidade	Garantia
	Suporte	Preço	Custo	Entrega	
S1	Consultoria	Benefícios	Mercado	Tecnologia	Retorno
	Fluxo	Fornecedor	Serviços	Suporte	Social
S2	Necessidade	Especificações	Preço	Qualidade	Garantia
	Suporte	Custo	Fornecedor	Entrega	

Apêndice A

Vetores numéricos das decisões

Este apêndice apresenta os dados coletados na dissertação de Bernardo Caldas [44] (apresentados no anexo A) em forma numeral.. As linhas representadas pela letra “D” correspondem as escolhas dos decisores e “P_D” as respectivas prioridades. Analogamente as linhas representadas pela letra “S” correspondem as escolhas dos subordinados e “P_S” as respectivas prioridades.

EMPRESA1										
D1	0,25	0,85	0,70	0,65	0,60	0,00	0,00	0,00	0,00	0,00
P _{D1}	1,00	0,90	0,80	0,70	0,60	0,00	0,00	0,00	0,00	0,00
S1	0,85	0,25	0,65	0,70	0,60	0,00	0,00	0,00	0,00	0,00
P _{S1}	1,00	0,90	0,80	0,70	0,60	0,00	0,00	0,00	0,00	0,00
S2	0,50	0,10	0,80	0,35	0,85	0,25	0,65	0,70	0,60	0,95
P _{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10
EMPRESA2										
D1	0,50	0,30	0,70	0,25	0,65	0,60	0,35	0,00	0,00	0,00
P _{D1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
S1	0,15	0,05	0,50	0,20	0,40	0,70	0,95	1,00	0,00	0,00
P _{S1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,00	0,00
S2	0,65	0,50	0,70	0,60	0,35	0,25	0,30	0,00	0,00	0,00
P _{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
EMPRESA3										
D1	0,70	0,25	0,65	0,40	0,95	0,20	0,35	0,60	0,05	0,00
P _{D1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,00
S1	0,05	0,65	0,25	0,35	0,70	0,60	0,95	0,20	0,40	0,00
P _{S1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,00
S2	0,50	0,10	0,30	0,15	0,05	0,25	0,95	0,55	0,80	0,40
P _{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10
EMPRESA4										
D1	0,70	0,35	0,65	0,95	0,60	0,40	0,20	0,00	0,00	0,00
P _{D1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
S1	0,05	0,35	0,95	0,15	0,70	0,65	0,60	0,40	0,00	0,00
P _{S1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,00	0,00
S2	0,35	0,40	0,95	0,70	0,65	0,60	0,20	0,00	0,00	0,00
P _{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
EMPRESA5										

D1	0,50	0,10	0,65	0,95	1,00	0,00	0,00	0,00	0,00	0,00
P_{D1}	1,00	0,90	0,80	0,70	0,60	0,00	0,00	0,00	0,00	0,00
S1	0,05	0,80	1,00	0,15	0,65	0,10	0,95	0,00	0,00	0,00
P_{S1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
S2	0,50	1,00	0,65	0,10	0,95	0,00	0,00	0,00	0,00	0,00
P_{S2}	1,00	0,90	0,80	0,70	0,60	0,00	0,00	0,00	0,00	0,00
EMPRESA6										
D1	0,25	0,50	0,20	0,15	0,70	0,75	0,95	0,05	0,65	0,60
P_{D1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10
S1	0,50	0,05	0,25	0,95	0,70	0,15	0,60	0,65	0,20	0,75
P_{S1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10
S2	0,10	0,50	0,05	0,85	0,25	0,70	0,40	0,95	0,15	0,80
P_{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10
EMPRESA7										
D1	0,50	0,15	0,35	0,65	0,60	0,80	0,05	0,00	0,00	0,00
P_{D1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
S1	0,80	0,15	0,30	0,35	0,95	0,45	0,50	0,20	0,25	0,40
P_{S1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10
S2	0,50	0,05	0,80	0,15	0,35	0,60	0,65	0,00	0,00	0,00
P_{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
EMPRESA8										
D1	0,35	0,45	0,55	0,70	0,25	0,50	0,20	0,40	0,65	0,60
P_{D1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10
S1	0,25	0,40	0,65	0,20	0,60	0,00	0,00	0,00	0,00	0,00
P_{S1}	1,00	0,90	0,80	0,70	0,60	0,00	0,00	0,00	0,00	0,00
S2	0,25	0,40	0,70	0,65	0,20	0,15	0,60	0,00	0,00	0,00
P_{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
EMPRESA9										
D1	0,45	1,00	0,50	0,15	0,80	0,35	0,75	0,00	0,00	0,00
P_{D1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
S1	0,50	0,45	0,15	0,80	1,00	0,75	0,35	0,00	0,00	0,00
P_{S1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,00	0,00	0,00
S2	0,50	0,45	0,05	0,15	0,8	1,00	0,25	0,20	0,65	0,00
P_{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,00
EMPRESA10										
D1	0,25	0,50	0,35	0,70	0,40	0,95	0,65	0,15	0,20	0,00
P_{D1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,00
S1	0,10	0,05	0,45	1,00	0,80	0,30	0,35	0,85	0,95	0,90
P_{S1}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10
S2	0,50	0,25	0,65	0,70	0,40	0,95	0,15	0,35	0,20	0,00
P_{S2}	1,00	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,00