

Concorrência, Redes de Petri e Teoria da Informação: Aspectos Comuns

Edmar Candeia Gurjão

Tese submetida à Coordenação dos Cursos de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Campina Grande como parte dos requisitos necessários para obtenção do grau de Doutor em ciências no domínio da Engenharia Elétrica.

Área de Concentração: Processamento da Informação

Angelo Perkusich, Doutor

Francisco Marcos de Assis, Doutor

Orientadores

Campina Grande, Paraíba, Brasil

©Edmar Candeia Gurjão, Junho de 2003



G979c Gurjão Edmar Candeia

**Concorrência, Redes de Petri e Teoria da Informação:
Aspectos Comuns/ Edmar Candeia Gurjão- Campina
Grande - PB: UFCCG, 2003.**

95p.: il.

Inclui Bibliografia

Tese (Doutorado) – UFCCG-CCT- Engenharia Elétrica

1 Redes de Petri 2. Capacidade 3. Informação

I. Título

CDU: 681.3.06

Dedicatória

A meu pai Heretiano Gurjão (*in memoriam*) e a minha mãe Ivanilda Candeia, que com simplicidade e sabedoria me ensinaram a trilhar os caminhos da vida.

Agradecimentos

- A minha esposa Sara pelo amor, pelos carinhos, incentivos e principalmente pela paciência comigo.
- A meus pais Heretiano (*in memoriam*) e Ivanilda (D. Nega) e meus irmãos Cléa, Elza, Eliza, Ednalva, Itamitã, Emir, Heretiano, Eraldo (*in memoriam*) e Edglay uma lista aparentemente grande, mas muito pequena se comparada a que seria necessária para apresentar as qualidades das pessoas citadas.
- Aos meus sobrinhos, a confirmação que de uma árvore boa e bem cuidada saem frutos muitos bons.
- Aos Professores Francisco Marcos e Angelo Perkusich que com grande sabedoria me guiaram nos caminhos difíceis de um trabalho de tese.
- Aos Professores do DEE-UFCG, em especial a Bruno Albert com quem tive o prazer de conviver durante esses anos de doutoramento e a Marcos Barbosa pelo incentivo a retomada do doutorado.
- Aos Funcionários do DEE-UFCG, em especial a Luiz Carlos e Angela que me deram o prazer de sua amizade, Pedro, Eleonora, Ronaldo, Rinaldo, Danilo e demais, sempre atenciosos e prestativos comigo.
- Aos Colegas de pós-graduação, que me proporcionaram momentos muito bons nesses anos de doutoramento.
- Aos demais amigos aqui representados por Ângelo Formiga, que de uma forma ou de outra sempre contribuíram para o meu crescimento pessoal e profissional.

Resumo

A Teoria da Informação tem grande sucesso quando aplicada a sistemas ponto-a-ponto, porém muitos problemas permanecem em aberto nos sistemas multi-ponto (sistemas de comunicação em rede), como por exemplo a determinação da região de capacidade (conjunto de taxas máximas simultâneas para um conjunto de transmissores) de certos canais. A transmissão da informação de uma fonte para um destino consiste na geração de eventos físicos no lado da fonte e na observação de eventos físicos no lado do destino, para sistemas em rede, tais eventos não podem ser totalmente ordenados. A ausência de ordenação entre eventos foi estudada no âmbito da ciência da computação por Carl A. Petri que introduziu precisamente a noção de "concorrência" entre eventos. Assim é natural esperar que os problemas para o cálculo de capacidade de redes de comunicação possam estar relacionados com os problemas da modelagem de sistemas concorrentes, sendo esta relação a principal motivação desta tese.

As redes de Petri são uma ferramenta bastante poderosa para a modelagem desses sistemas capturando os conceitos de estado distribuído, as relações de concorrência e causalidade entre eventos. Nesta tese inicia-se o uso de redes de Petri para a análise da capacidade de sistemas de comunicação multi-ponto. Neste sentido, é feita uma extensão de uma medida de complexidade definida para as rede de Petri, denominada de compressibilidade. Mostra-se que a compressibilidade de uma rede pode ser calculada a partir de uma matriz de adjacências associada ao seu grafo de cobertura. Também são apresentadas algumas propriedades dessa medida. Usando a compressibilidade, são obtidas algumas conclusões a respeito da representação e do comportamento de sistemas concorrentes.

São apresentados exemplos de sistemas não-ruídos concorrentes que geram símbolos obedecendo restrições semelhantes às restrições de um sistema de gravação magnética. Apresenta-se um método para o cálculo da capacidade desses sistemas por meio da modelagem em redes de Petri e de uma função de rotulação convenientemente definida que associa eventos físicos com informação.

Abstract

Information theory is a successfully discipline concerning point-to-point systems but many problems in networked communication systems remain open as for example the capacity region of some multi-user channels are unknown. Considering the very basic fact that trasmission of information from a source point to a destination point is due by choosing physical events on the transmitter side in order to cause physical events on the receiver side, it is observed that in case of networked systems in general it is impossible to obtain a total ordering of such that events. In other hand lack of ordering of events in distributed systems was explored by Carl A. Petri that has introduced the precise notion of "concurrency" between events in the circle of the computer science. The today well known Petri nets efficiently model aspects os concurrence and causality between events. In view of comments above it is natural ask about relationship between problems of capacity of networked communication systems and modelling of concurrent systems. This is the main motivation of this thesis.

In this thesis it started the use of Petri nets in the analysis of the capacity of multi-point systems. To this end, a previous defined measure of complexity of Petri nets named compressibility is extended. It is shown that the compressibility of a Petri net can obtained from an adjacency matrix associated with its coverability graph. Some of the properties of this measure are presented. From compressibility definition, some conclusions related to the representation and the behavior of concurrent systems are presented.

Examples of concurrent noiseless systems generating sequences of symbols that obey some constraints similar to those found in magnetic recording systems. It is shown how to calculate the capacity of such that systems by means Petri nets modelling and a handy labeling function that associates physical events to information.

Lista de Símbolos e Abreviaturas

- \wedge - Operação AND lógico;
- \neg - Operação NOT lógico;
- \Rightarrow - Implica;
- \prec - Relação de dependência causal;
- \Leftrightarrow - Se e somente se;
- λ - Maior autovalor de uma matriz;
- A** - Símbolos em negrito representam matrizes;
- BSC* - Canal Binário Simétrico (do inglês *Binary Symmetric Channel*);
- c* - Velocidade de luz, $c = 300000 \text{ km/segundo}$;
- C* - Capacidade;
- co* - Relação de Concorrência;
- CS* - Estrutura de Concorrência (do inglês, *Concurrency structure*);
- CPN* - Rede de Petri Colorida (do inglês, *Coloured Petri Net*);
- DEP* - Densidade Espectral de Potência;
- DNC* - Canal Discreto sem Ruído (do inglês, *discrete noise channel*);
- DNC - 2D* - Canal Discreto sem Ruído Bi-dimensional;
- FSTD* - Diagrama de Transição de Estados Finitos (do inglês, *finite state transition diagram*);
- H* - Entropia;
- H_{PN}* - Compressibilidade;
- $I(X;Y)$ - Informação Mútua entre as variáveis aleatórias *X* e *Y*;

li – Relação de Causalidade;

MAC – Canal de Acesso Múltiplo (do inglês, *multiple access channel*);

NRZ – Não retorno ao zero (do inglês, *non-return-to-zero*;

NRZI – Não retorno ao zero invertido (do inglês, *non-return to zero inverted*);

SPO – Ordem parcial estrita (do inglês, *strict partial order*);

Lista de Figuras

2.1	Diagrama de transição estados finitos com três estados.	10
3.1	Eventos no sistema de coordenadas espaço \times tempo.	16
3.2	Exemplo de rede de Petri.	20
3.3	Modelo em rede de Petri para o sistema com dois usuários.	23
3.4	Modelo em rede de Petri para o sistema com dois usuários após a sequência de transições t_2t_3	24
3.5	Marcação após o disparo de t_2	25
3.6	Marcação após o disparo de t_4	25
3.7	Grafo de alcançabilidade para a rede de Petri da Figura 3.3.	26
3.8	Uma rede de Petri ilimitada e seu grafo de cobertura.	27
3.9	Modelo em rede de Petri Colorida para o sistema com dois usuários. . .	30
3.10	Modelo para o sistema com dois usuários e um recurso em comum. . . .	32
4.1	Representação de dois usuário escrevendo independentemente em um mesmo papel.	40
4.2	Modelo em rede de Petri para $ T $ eventos concorrentes	41
4.3	Grafo de alcançabilidade para o modelo da Figura 4.2	41
4.4	Modelo em rede de Petri para o sistema de dois canais em paralelo com restrição.	46
4.5	Grafo de alcançabilidade do Modelo em rede de Petri para o sistema de dois canais em paralelo com restrição.	46

4.6	Modelo em rede de Petri para o sistema de comunicação por cores. . . .	48
4.7	Canal com acesso múltiplo.	51
4.8	Região de capacidade de um canal com acesso múltiplo aditivo.	51
4.9	Modelo em para o canal MAA.	52
4.10	Grafo de alcançabilidade para o canal MAA.	53
4.11	Grafo de alcançabilidade reduzido.	53
4.12	Grafo de alcançabilidade reduzido considerando somente as transmissões do primeiro usuário.	54
4.13	Modelo em rede de Petri Colorida para o código $(d, k; n)$	56
4.14	Grafo de alcançabilidade para o modelo em rede de Petri para o código $(1, 3; 2)$	58
4.15	Grafo de ocorrência para o modelo em rede de Petri do código $(1, 2; 1, 3; 2)$	59
4.16	Modelo para o código $(c, j; d, k; n)$	60
A.1	Diagrama de estados finitos para as seqüências $M(d, k)$	67
A.2	FSTD para as probabilidades de transição.	68
A.3	Canal com apagamento cruzado.	73
A.4	Limitante inferior da Capacidade em função de α para $d = 0$ e $k = 3$ a) $M = 2$, b) $M = 3$, c) $M = 4$ e d) $M = 5$	75
A.5	Limitante inferior da Capacidade em função de α para $d = 1$ e $k = 3$ a) $M = 2$, b) $M = 3$, c) $M = 4$ e d) $M = 5$	76
A.6	Aplicação da partição de estados.	81
A.7	FSTD para $M = 2$, $(d, k) = (1, 2)$	82
A.8	DEP para $M = 2$ e $(d, k) = (1, 2)$	83
A.9	Codificação NRZI para o caso binário.	83
A.10	Codificação NRZI para o caso binário.	84
A.11	DEP para $(d, k) = (1, 3)$ e variando o valor de M	84
A.12	Esquema NRZI mod M	84
A.13	DEP para o esquema NRZI mod M	85

B.1 Aplicação da construção por subconjuntos.	89
---	----

Lista de Tabelas

2.1	Exemplo de codificação unívoca de símbolos de uma fonte.	6
4.1	Função de rotulação para o grafo de alcançabilidade para os <i>passos</i> do modelo da Figura 4.4	47
4.2	Função de rotulação para o grafo de alcançabilidade apresentado na Figura 4.3	49
4.3	Outra função de rotulação para o grafo de alcançabilidade apresentado na Figura 4.3	50
4.4	Passos viáveis do modelo apresentado na Figura 4.9.	51
4.5	Marcações para o código $(1, 3; 2)$	58
4.6	Valores de taxas para alguns códigos $(c, j; d, k; n)$, que não foram calculados em [1].	61

Conteúdo

1	Introdução	1
2	Medidas de Informação	4
2.1	Entropia	4
2.2	Capacidade	7
2.2.1	Capacidade de Canais sem Ruído	8
2.2.2	Cálculo da Capacidade usando Matrizes de Adjacências	9
2.2.3	Capacidade de Canais com Ruído	11
2.3	Capacidade em Sistemas Multi-ponto	12
2.4	Conclusões	13
3	Redes de Petri e Outros Modelos de Concorrência	15
3.1	Concorrência	16
3.2	Redes de Petri	19
3.2.1	Regra de disparo	22
3.3	Redes de Petri Coloridas	28
3.4	Outros Modelos para Sistemas Concorrentes	31
3.5	Conclusões	33
4	Compressibilidade de Redes de Petri e Capacidade de Sistemas Con-	
	correntes	34
4.1	Introdução	34

4.2	Compressibilidade das Redes de Petri	35
4.2.1	Calculando a Compressibilidade	38
4.3	Capacidade de Sistemas Concorrentes com Restrição	43
4.3.1	Construção dos Modelos Usando os Passos Viáveis	48
4.3.2	Aplicações	53
4.4	Conclusões	61
5	Conclusões	62
A Capacidade de Canais M-ários com Restrição (d, k) nas Seqüências de		
	Entrada	65
A.1	Introdução	66
A.2	Códigos $M(d, k)$	67
A.3	Probabilidades para seqüências M -árias	68
A.4	Canal com Cruzamento	72
A.5	Limitante Inferior para a Capacidade	72
A.5.1	Exemplos de Limitantes	74
A.6	Análise Espectral de Cadeias de Markov	75
A.6.1	Densidade Espectral de Funções de uma Cadeia de Markov	76
A.7	Aplicação aos códigos $M(d, k)$	80
A.8	Conclusões	85
A.9	Cálculo da Informação Mútua $I(S_2; Z S_1)$	86
B Construção por Subconjuntos (<i>Subset Construction</i>)		
B.0.1	Observações	89

Capítulo 1

Introdução

Desde o seu lançamento com o trabalho de Shannon [2], a Teoria da Informação vem obtendo grande sucesso nas mais diversas áreas do conhecimento, dentre elas destacam-se as Telecomunicações, a Física e a Ciência da Computação [3]. Essa teoria facilita a compreensão de diversos problemas, contribui para o desenvolvimento de sistemas e fornece limitantes fundamentais para a transmissão da informação.

Para os sistemas ponto-a-ponto, uma fonte e um destino, Shannon definiu medidas objetivas tanto para a quantidade de informação gerada pela fonte (entropia) como para a quantidade máxima de informação que pode ser transmitida (capacidade), com probabilidade de erro assintoticamente nula, para o destino através de um meio de transmissão. Essas medidas e os métodos para calculá-las já estão bem definidos.

Nos sistemas de comunicação em rede, que são sistemas multi-ponto com várias fontes e destinos, a teoria da informação ainda não obteve o mesmo sucesso, pois as medidas de informação para esses sistemas ainda não estão bem estabelecidas.

A aplicação da teoria da informação em sistemas multiponto teve seu início na década de 60 com outro trabalho de Shannon [4]. Nesse trabalho, foram estabelecidos os limitantes inferior e superior para a quantidade máxima de informação que podem ser enviada simultaneamente pelas diversas fonte através de um determinado meio de transmissão. Muito tem sido realizado nessa área [5], [6] e [7], mas ainda hoje observa-

se que existem muitas questões em aberto [8]. Como exemplo, os limitantes definidos por Shannon em [4] ainda são os únicos resultados, em termos de taxas de transmissão máximas, conhecidos para o meio ali tratado, e para outros cenários nada pode ser afirmado, como nos canais com interferência [7]. Portanto, observa-se que os métodos atualmente usados para analisar esses canais não fornecem as respostas desejadas, e assim novos métodos de análise devem ser procurados, é nesse sentido que esta tese se posiciona.

Inicialmente observa-se que os os sistemas multi-ponto são essencialmente sistemas distribuídos, e para entender os fenômenos que neles ocorrem, considere que a transmissão de informação de um ponto do espaço (espaço-tempo) para outro possa ser vista como a *escolha de eventos físicos* no lado do transmissor para *causar eventos físicos* no lado do receptor. Em sistemas de comunicação ponto-a-ponto (um transmissor e um receptor) a ordem em que os eventos são escolhidos no transmissor ou que são causados no receptor pode ser completamente descrita, e assim diz-se que esses eventos estão completamente ordenados. Entretanto, para sistemas de comunicação em redes (vários transmissores vários receptores) devido à impossibilidade de sincronização ou a distância relativa entre os transmissores, ou entre esses e os receptores, esse ordenamento nem sempre é possível, nesse caso diz-se que os eventos estão parcialmente ordenados. Esta impossibilidade de ordenamento de eventos define a relação de *concorrência* entre esses eventos [9].

É possível modelar esses sistemas usando diagrama de estados finitos, mas mostra-se que a complexidade desses modelos pode aumentar exponencialmente com o número de estados do sistema [10]. Para solucionar esse problema, outras formas de modelar foram propostas. Uma dessas formas é usando redes de Petri que são uma ferramenta matemática muito eficiente para tratar com sistemas concorrentes. Essas redes tem sido usadas em muitas áreas da ciência da computação e sistemas a eventos discretos incluindo engenharia de software, banco de dados, sistemas de informação, protocolos de comunicação e redes de computadores [11, 10].

A teoria de redes de Petri permite modelar sistemas que apresentam concorrência,

conflito, paralelismo e distribuição, portanto pode-se investigar a aplicação da teoria de Petri no contexto da teoria da informação para redes, esta é a motivação principal desse trabalho. Por outro lado, também é possível questionar como os resultados da teoria da informação podem ser aplicados no contexto da teoria de redes de Petri para prover novos resultados e metodologias.

Neste trabalho, propõe-se um método para o cálculo da capacidade de sistemas que apresentam concorrência e cujas seqüências de símbolos devem obedecer a restrições. O método proposto faz uso das redes de Petri e da extensão de uma medida de complexidade previamente definida na literatura para essas redes [12]. A extensão da definição de complexidade para redes de Petri constitui a primeira contribuição deste trabalho. Usando o modelo do sistema em rede de Petri e uma função de rotulação dos eventos do sistema, que são obtidos do modelo, tem-se a segunda contribuição que consiste na definição de um sistema concorrente com restrição e de um método para o cálculo da capacidade desses sistemas. Através desse método, é estabelecida uma relação entre a capacidade e a complexidade do modelo em rede de Petri de um sistema concorrente com restrição. Em seguida, são apresentados exemplos de aplicação do método proposto.

No texto desta tese tem-se no Capítulo 2 uma introdução as medidas de informação definidas por Shannon, no Capítulo 3 uma introdução à Teoria Concorrência, a descrição das redes de Petri e algumas de suas variantes. Em seguida no Capítulo 4, define-se a versão estendida da medida de complexidade para as redes de Petri, e o método de cálculo da capacidade dos sistemas concorrentes com restrição junto com alguns exemplos. No Capítulo 5 são apresentadas as conclusões e as perspectivas de trabalhos futuros. O texto conta com dois apêndices. No Apêndice A é feito o cálculo de um limitante para a capacidade de canais cujas seqüências de entrada tem restrição e é obtida a densidade espectral de potência dessas seqüências. No Apêndice B tem-se um método de operação em grafos rotulados que é usado no Capítulo 4.

Capítulo 2

Medidas de Informação

Neste capítulo são apresentadas algumas medidas de informação propostas por Shannon. Inicialmente, define-se a entropia como a medida da quantidade de informação gerada por uma fonte ou por um sistema. Em seguida, apresenta-se a medida da máxima quantidade de informação que pode ser transmitida através de um meio, essa medida é denominada de capacidade do canal. Para essa última medida dois casos são considerados. No primeiro considera-se o canal sem ruído e no segundo a presença do ruído. Para cada situação é descrita uma maneira de calcular a capacidade. Por simplicidade, trataremos de fontes e sistemas discretos e sem memória.

2.1 Entropia

Os sistemas de comunicações são constituídos basicamente de uma fonte, um meio de transmissão e um destino. A fonte gera símbolos pertencentes a um alfabeto, e que obedecem a uma distribuição de probabilidades. Os símbolos (ou blocos deles) são enviados através de um meio, a partir daqui denominado canal de comunicação ou simplesmente canal, para um destino. No destino, observando a saída do canal, tem-se como objetivo obter uma estimativa confiável da informação enviada. A comunicação pode ser de um transmissor para um receptor como nos canais ponto-a-ponto, ou de

vários transmissores para diversos receptores como nos canais multi-ponto (também chamados de sistemas de comunicação em rede).

Shannon [2] definiu medidas objetivas tanto para a quantidade de informação gerada pela fonte, como para a quantidade máxima de informação que pode ser transmitida de forma confiável através de um canal.

No caso da fonte, cujos símbolos pertencem a um conjunto \mathcal{X} , baseado na distribuição de probabilidades dos símbolos (p_i) foi definida a medida de informação denominada entropia (H)

$$H(\mathbf{p}) = - \sum_{i=1}^{|\mathcal{X}|} p_i \log p_i \quad (2.1)$$

sendo $\mathbf{p} = (p_1, p_2, \dots, p_{|\mathcal{X}|})$ o vetor de probabilidade dos $|\mathcal{X}|$ símbolos da fonte e o logaritmo na base 2¹. A entropia é expressa em bits por símbolo. Ela fornece uma idéia da complexidade da fonte, pois indica o número mínimo de bits que pode ser usado para representar cada símbolo univocamente [2].

Por exemplo, considere uma fonte que gera símbolos pertencentes ao alfabeto $\mathcal{X} = \{A, B, C, D\}$ com distribuição $p_A = 1/2$, $p_B = 1/4$, $p_C = 1/8$ e $p_D = 1/8$. A entropia dessa fonte é

$$\begin{aligned} H(\mathbf{p}) &= - \sum_{i=1}^4 p_i \log p_i, \\ &= 1,75 \text{ bits por símbolo.} \end{aligned} \quad (2.2)$$

O valor da entropia dessa fonte representa o número mínimo de bits necessário para representar univocamente cada um dos símbolos emitidos por ela, por exemplo observe a Tabela 2.1. O código COD_1 usa na média 1,75 bits por símbolo e cada palavra código representa univocamente cada um dos símbolos da fonte. Outra possibilidade é usar palavras código de tamanho fixo como o código COD_2 que usa 2 bits por símbolo, observe que nesse caso os símbolos da fonte também são representados univocamente, porém na média são usados mais bits para cada símbolo do que em COD_1 . Observando

¹Neste trabalho, a menos que seja explicitamente especificada outra base, todos os logaritmos são tomados na base 2

a entropia como uma medida de incerteza, o valor obtido indica o número *mínimo* de perguntas binárias para determinar qual símbolo foi emitido [13].

x	p_x	COD_1	COD_2
A	$1/2$	0	00
B	$1/4$	10	01
C	$1/8$	110	10
D	$1/8$	111	11

Tabela 2.1: Exemplo de codificação unívoca de símbolos de uma fonte.

Seja a mesma fonte gerando símbolos com probabilidades $p_A = 1$, $p_B = 0$, $p_C = 0$ e $p_D = 0$. Nesse caso $H(\mathbf{p}) = 0$, observando esse valor como uma medida de incerteza dos símbolos gerados pela fonte, vê-se que não há incerteza alguma com relação aos símbolos de saída.

A medida de entropia definida acima trata com apenas uma variável aleatória. Porém, quando se tem uma seqüência de n variáveis aleatórias pode-se questionar como a entropia dessa seqüência cresce com n . Para tanto define-se a taxa de entropia como

$$H(\mathcal{X}) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n) \quad (2.3)$$

quando o limite existe. Por exemplo, para variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) tem-se que

$$H(\mathcal{X}) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n) = \lim_{n \rightarrow \infty} \frac{nH(X_1)}{n} = H(X_1).$$

Essa medida será usada no Apêndice A para tratar de fontes cujas seqüências de símbolos devem obedecer a restrições.

A seguir a medida da quantidade de informação que pode ser transmitida por um canal é apresentada.

2.2 Capacidade

O canal de comunicação é o meio pelo qual os símbolos serão transmitidos. Podemos considerar a transmissão de um ponto no espaço para outro, como na transmissão de dados entre duas cidades, ou de um ponto no tempo para outro como nos sistemas de gravação [14].

Em ambos os casos, o canal tem como entrada uma mensagem cujos símbolos pertencem a um alfabeto finito \mathcal{X} . Devido ao ruído, que pode alterar as mensagens enviadas, a mensagem observada na saída do canal não é necessariamente a mesma que foi colocada na entrada. A capacidade (C), medida em bits por uso do canal, que representa a taxa máxima de transmissão de símbolos pelo canal sob uma probabilidade de erro (ϵ) assintoticamente nula.

O Teorema a seguir, demonstrado por Shannon em [2], relaciona a entropia da fonte com a capacidade do canal de comunicações.

Teorema 2.1 *Seja uma fonte com entropia H (bits por símbolo) e um canal com capacidade C (bits por uso do canal). É possível codificar a saída da fonte para transmitir em média $\frac{C}{H} - \epsilon$ símbolos por uso do canal sendo ϵ um valor arbitrariamente pequeno. Não é possível transmitir, de forma confiável, em média, a taxas maiores que $\frac{C}{H}$.*

Observe que taxas maiores que C/H podem ser usadas, porém, nesse caso, não há garantia alguma que a probabilidade de erro seja assintoticamente nula. O valor da capacidade é um dos parâmetros de maior importância no projeto de sistemas ou códigos para transmitir em um determinado canal. Por exemplo, considere uma fonte que gera $|\mathcal{X}|$ símbolos equiprováveis, a entropia dessa fonte é $\log |\mathcal{X}|$ bits por símbolo. Se a informação dessa fonte for transmitida através de um canal que tenha capacidade de $C_1 = \log |\mathcal{X}|$ bits por uso, a cada transmissão (uso do canal) toda a informação gerada pela fonte será passada pelo canal. Em outro canal com capacidade $C_2 = \log |\mathcal{X}| / 4$ bits por uso, serão necessárias no mínimo 4 transmissões para que a informação gerada pela fonte possa chegar ao destino. O valor da capacidade

é de grande importância para o conhecimento do canal e de suas limitações, portanto pode ser usado como parâmetro básico no projeto dos sistemas e dos códigos que serão usados na transmissão através desse canal.

A determinação da capacidade de um canal de comunicação pode ser feita considerando dois casos. No primeiro considera-se o canal sem ruído e no segundo considera-se a presença do ruído. Para cada caso existem métodos de cálculo da capacidade. A seguir na Seção 2.2.1 o primeiro caso é apresentado e na Seção 2.2.3 tem-se o segundo caso.

2.2.1 Capacidade de Canais sem Ruído

Neste trabalho trataremos de canais e sistemas discretos.

Definição 2.1 *Um canal discreto sem ruído (DNC, do inglês discrete noiseless channel) associado a um conjunto de seqüências \mathcal{S} , é aquele que admite a transmissão somente de seqüências pertencentes a \mathcal{S} , sendo esta transmissão livre de erros.*

A capacidade (C) de um DNC foi definida como [2]

$$C = \lim_{\tau \rightarrow \infty} \frac{\log_2 N(\tau)}{\tau}, \quad (2.4)$$

sendo $N(\tau)$ o número de seqüências de comprimento τ em \mathcal{S} e C a taxa máxima de transmissão da informação através do DNC.

O canal acima descrito é dito ser unidimensional pois ele permite a transmissão de um símbolos por vez, a seguir é definido o canal sem ruído que admite transmissões de símbolos em paralelo, denominados de vetores.

Definição 2.2 *Um canal discreto sem ruído bidimensional (DNC-2D) associado a um conjunto de seqüências \mathcal{S} , é aquele que admite a transmissão somente de seqüências de vetores de comprimento m , denominados vetores de canal, cujos elementos pertencem a \mathcal{S} e satisfazem a restrições nas seqüências desses vetores, sendo esta transmissão livre de erros.*

No canal DNC-2D denomina-se um vetor com todos os componentes zero por vetor zero de canal. Os vetores de canal podem ser por exemplo, os m símbolos lidos/gravados em paralelo em um sistema que permita a gravação em várias trilhas.

A capacidade por componente para o canal DNC-2D, que tem vetores de canal com m componentes é definida como

$$C = \frac{1}{m} \lim_{\tau \rightarrow \infty} \frac{\log_2 N_m(\tau)}{\tau}, \quad (2.5)$$

sendo $N_m(\tau)$ o número de seqüências de vetores de canal comprimento τ . A multiplicação do limite por $(1/m)$ fornece a capacidade por componente, esse cálculo tem por objetivo fornecer uma base de comparação com o caso unidimensional.

2.2.2 Cálculo da Capacidade usando Matrizes de Adjacências

Nesta seção, é apresentado um método para o cálculo da capacidade de canais sem ruído [2].

Um diagrama de transições de estados finitos (*FSTD* do inglês, *Finite State Transition Diagram*) consiste de um grafo cujos vértices representam os estados do sistema, e um arco entre dois estados representa uma transição entre eles.

Dado um diagrama de estados finitos, é possível associar a ele uma matriz denominada matriz de adjacências definida a seguir.

Definição 2.3 *A matriz de adjacências de um diagrama de transição de estados finitos com n estados, denotada por A é uma matriz $n \times n$ sendo o elemento (i, j) dessa matriz o número de arcos nesse diagrama do estado i para o estado j .*

Desse modo, a matriz de adjacências contém a informação das quantidades de caminhos de tamanho um entre dois estados quaisquer, porém ela não mostra qual símbolo foi emitido nessa troca. Por exemplo, seja o diagrama de transição de estados finitos apresentado na Figura 2.1, sua matriz de adjacência é apresentada na Equação 2.6. Observe que as linhas e as colunas de A são indexadas pelos estados do diagrama, nesse caso S_0 , S_1 e S_2 .

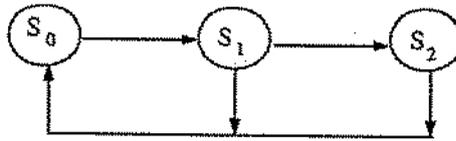


Figura 2.1: Diagrama de transição estados finitos com três estados.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad (2.6)$$

A soma dos elementos da matriz A é igual a quantidade de caminhos de comprimento um no FSTD correspondente. Sendo A^l a multiplicação dessa matriz por ela mesma l vezes, os elementos a_{ij} da matriz resultante representam a quantidade de caminhos de comprimento l entre os estados i e j . E assim, a soma de todos os elementos de uma matriz A^l é igual a quantidade de caminhos de comprimento l no diagrama a que essa matriz está associada.

Assim, a matriz de adjacência fornece a quantidade de caminhos de comprimento τ , ($N(\tau)$), através da soma dos elementos de A^τ . Portanto, é possível usá-la para calcular a capacidade do sistema que ela representa. Mostra-se [2], [15], que a capacidade de um sistema com matriz de adjacências A , é igual ao logaritmo do maior autovalor dessa matriz. A teoria de Perron-Frobenius [15, p. 106] garante que as matrizes de adjacências tem somente um autovalor real positivo máximo.

Caso os arcos do FSTD sejam rotulados, o método do cálculo da capacidade usando a matriz de adjacências somente é válido se em cada estado desse diagrama os arcos que partem forem rotulados com símbolos distintos. Um diagrama desse tipo é dito ser resolvido a direita (do inglês, *right-resolving*) [15]. Diagramas resolvidos a direita são ditos determinísticos, pois dados uma palavra w e um estado I desse diagrama existe no máximo um caminho que inicia em I e é formado pelos símbolos de w . Por simplicidade, já que um FSTD é um grafo, nessa tese será usado o termo grafo determinístico para

um diagrama resolvido a direita. Caso o diagrama não seja resolvido a direita, para que seja possível aplicar o método da matriz de adjacência para o cálculo da capacidade, deve ser aplicado um algoritmo que gera um novo diagrama que representa as mesmas seqüências de símbolos mas é resolvido a direita. Um desses algoritmos é a construção por subconjuntos apresentado no Apêndice B.

Uma das dificuldades para obter a capacidade usando o procedimento aqui descrito é a obtenção do FSTD e conseqüentemente a matriz de adjacência. Em alguns casos, se o número de estados for pequeno, é possível obtê-la usando uma busca exaustiva pela observação para cada um dos estados quais as adjacências. Entretanto, a medida que a quantidade de estados aumenta esse procedimento torna-se impraticável. Portanto, outros métodos que possam fornecer essa o FSTD ou a própria matriz devem ser buscados. Uma das soluções para esse problema é obter métodos recursivos que forneçam essas matrizes, outra possibilidade é usar alguma modelagem que forneça de forma sistemática os estados e as relações entre eles. Como exemplo desse último método, Sindhuhsyana [16] em sua dissertação de mestrado propôs um método que parte de uma lista de seqüências de símbolos não permitidos pelo sistema e obtém um grafo que representai-o.

O cálculo da capacidade através da matriz de adjacências de um canal é válido somente quando o canal é sem ruído, na presença de ruído a análise da capacidade de um canal muda conforme apresentado na próxima seção.

2.2.3 Capacidade de Canais com Ruído

O ruído no canal de comunicação pode modificar algumas características da mensagem transmitida, e portanto a mensagem recebida não é necessariamente a enviada. A mensagem transmitida ($X \in \mathcal{X}$) quando passa pelo canal perde uma dada quantidade de informação e na saída do canal tem-se a mensagem recebida ($Y \in \mathcal{Y}$). Vale observar que dependendo do canal os alfabetos de entrada e saída podem ser diferentes. A perda de informação pode ser medida usando a entropia de Y condicionada a X , com

distribuição conjunta $p(x, y)$, definida como

$$H(X | Y) = - \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log p(y | x). \quad (2.7)$$

Usando a entropia de Y e a entropia condicionada $H(Y | X)$ têm-se a informação mútua entre a entrada e saída, dada por

$$I(X; Y) = H(Y) - H(Y | X).$$

A capacidade do canal é definida como o máximo da taxa acima.

Definição 2.4 (Capacidade na presença de ruído) *A capacidade C de um canal discreto sem memória ruidoso, com entrada $X \in \mathcal{X}$ com distribuição p , e saída $Y \in \mathcal{Y}$ é a máxima taxa de transmissão é dada por*

$$C = \max_p I(X; Y)$$

sendo o máximo com relação a todas as possíveis distribuições de entrada p .

Se o canal não tem ruído e a fonte tem alfabeto $|\mathcal{X}|$, então $Y = X$ e $H(Y | X) = 0$ então $I(X; Y) = H(X)$. Nesse caso retorna-se ao canal sem ruído, e o método da seção anterior pode ser usado. Como o máximo da entropia é $H(X) = \log |\mathcal{X}|$ para uma distribuição p equi-provável tem-se que $C = \log |\mathcal{X}|$, esse é o máximo valor da capacidade. No outro extremo, se o ruído é tal que a entrada e saída se tornam independentes $H(Y | X) = H(Y)$ e $I(X; Y) = 0$ então $C = 0$, não é possível transmitir nesse canal.

2.3 Capacidade em Sistemas Multi-ponto

No caso de sistemas ponto-a-ponto, a capacidade é suficiente para representar a máxima taxa de informação que pode ser transmitida, pois têm-se apenas um transmissor e um receptor. Para sistemas multi-ponto o conceito de capacidade tem que ser estendido

para uma região de capacidades. Uma região de capacidade define os conjuntos de taxas máximas que os usuários podem transmitir simultaneamente.

A Teoria da Informação multi-usuário teve início com o trabalho de Shannon [4], no qual foram introduzidos os canais bidirecionais (*two-way*). Nesses canais dois usuários trocam informações simultaneamente usando um canal comum e as informações em um sentido interferem nas enviadas no sentido oposto. Shannon definiu os limitantes inferior e superior para a região de capacidade e ainda hoje somente se conhece esses limitantes. Alguns códigos definidos para esses canais atingiram taxas acima do limitante inferior [17]. Várias técnicas tem sido usadas na busca da região exata, por exemplo o uso de fractais [6], mas essa região ainda não foi encontrada.

Existem outros canais multi-usuários como os canais de Acesso Múltiplo, nos quais vários usuários enviam informação pelo mesmo canal para um destino [5] e o canal de difusão (*Broadcast*) no qual uma fonte transmite informações para vários usuários por um mesmo canal [5].

É importante observar que as técnicas usada atualmente para determinar a região de capacidade de sistemas multi-ponto não foram suficientes para determinar essa região, e assim deve-se procurar outras metodologias que possibilitem a análise desses canais para que possam ser usados nos cálculos de suas regiões de capacidade.

2.4 Conclusões

Neste capítulo foram apresentadas algumas medidas de informação definidas por Shannon. No caso da entropia de uma fonte, foi mostrado que ela pode ser obtida pela taxa de crescimento da quantidade de seqüências distintas que essa fonte produz. No caso do canal, foi apresentada a medida da quantidade de informação que ele pode transmitir e foram considerados dois casos, a ausência e presença do ruído. Para cada caso foi apresentada a definição da capacidade.

Tanto na definição de entropia quanto da capacidade pressupõe-se um ordenamento dos símbolos emitidos pela fonte ou enviado/transmitidos pelo transmissor/receptor.

No próximo capítulo será mostrado que nem todos os sistemas tem a característica que os símbolos sejam ordenados, e para tratar com esse tipo de sistema será definida a concorrência e algumas ferramentas para modelar os sistemas que apresentam esse fenômeno. Usando uma dessas ferramentas, no caso as redes de Petri, será mostrado que é possível calcular as medidas de informação para sistemas concorrentes usando essas redes como base.

Capítulo 3

Redes de Petri e Outros Modelos de Concorrência

As medidas de entropia e capacidade requerem um conhecimento prévio do comportamento (em termos de distribuições de probabilidades) da fonte e do canal respectivamente. A análise do comportamento de um sistema pode ser feita através da observação do próprio sistema ou de um modelo para ele, o que normalmente é preferível principalmente numa fase de projeto. Para isto, faz-se necessário uma ferramenta de modelagem que seja capaz de capturar os fenômenos que ocorrem no sistema para que seja possível representá-lo o mais fielmente possível. Neste capítulo, é feita uma introdução a uma ferramenta de modelagem chamada redes de Petri. Essas redes tem dentre outras características, a de serem apropriadas para modelar sistemas que apresentam concorrência.

Inicialmente, partindo do princípio que a cada ocorrência de um evento pode-se associar um ponto no sistema de coordenadas espaço×tempo, é feita uma introdução à concorrência e mostra-se como esse fenômeno é modelado pelas redes de Petri. Em seguida, as redes de Petri são definidas e algumas de suas características são apresentadas. Outras formas de modelar a concorrência são descritas e finalmente são apresentadas algumas conclusões.

3.1 Concorrência

Considere a execução de um processo qualquer, descrita com o nível de detalhes desejado. Os elementos x que representam as ocorrências do processo podem ser identificados por coordenadas $x = (r(x), t(x))$ em que $r(x)$ refere-se as coordenadas do espaço e $t(x)$ as do tempo, esses elementos são agrupados em um conjunto X . Se esse conjunto é ordenado por uma relação binária \prec que indica dependência causal, então $a \prec b$ se e somente se um sinal físico inicia em a e termina em b , sendo $a, b \in X$ [18].

A velocidade dos sinais físicos está limitada pela velocidade de luz c . Se um sistema de coordenadas espaço \times tempo é construído (Figura 3.1) e são marcadas as ocorrências de x e y , tem-se Δr a distância espacial entre x e y , e Δt a diferença de tempo entre suas ocorrências, então

$$x \prec y \implies (\Delta r)^2 \leq c^2(\Delta t)^2 \wedge \Delta t > 0. \quad (3.1)$$

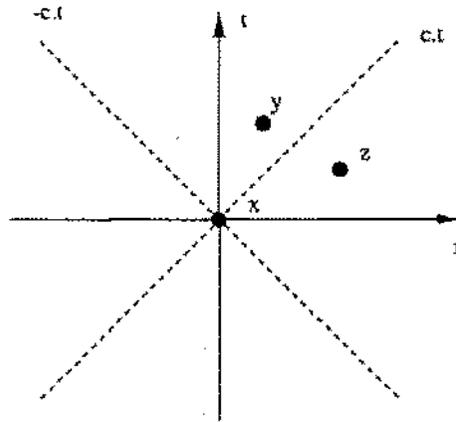


Figura 3.1: Eventos no sistema de coordenadas espaço \times tempo.

Na Figura 3.1 as curvas $c \cdot t$ e $-c \cdot t$ correspondem as linhas que um sinal que parte de x com a maior velocidade possível (velocidade da luz c) descreve. Os eventos dentro da região definida por essas linhas (denominada de cone de luz), por exemplo y , podem receber os sinais enviados de x obedecendo a Equação 3.1. Para os eventos fora dessa

região, por exemplo z , não é possível satisfazer a relação apresentada na Equação 3.1 e portanto não pode haver troca de sinalização entre esses eventos e x .

Assim, é possível definir uma ordem parcial estrita (SPO do inglês, *strict partial order*) $SPO(X, \prec)$.

Definição 3.1 (Ordem parcial estrita)

$$\begin{aligned} SPO(X, \prec) : \forall x, y : x \prec y \Rightarrow x \in X \wedge y \in X \\ \wedge \quad \forall x \in X : \neg(x \prec x) \\ \wedge \quad \forall x, y, w \in X : x \prec y \wedge y \prec w \Rightarrow x \prec w. \end{aligned}$$

Numa ordem parcial estrita é possível definir uma relação de “desordem” como $\neg(x \prec y)$ e $\neg(y \prec x)$. Essa relação foi definida por Carl A. Petri e é chamada de concorrência (*co*) [18], ou seja

$$x \text{ co } y \Leftrightarrow x, y \in X \wedge \neg(x \prec y) \wedge \neg(y \prec x). \quad (3.2)$$

Então, a ordem parcial estrita (X, \prec) pode ser vista como uma representação da história das ocorrências de um processo. Se X é totalmente ordenado¹, então $x \text{ co } y \Leftrightarrow x = y$ e a concorrência fica confinada a identidade [9].

Fisicamente, quando da ocorrência de dois eventos a e b diz-se que eles são concorrentes se não for possível satisfazer a relação apresentada na Equação 3.1, pois nesse caso não é possível qualquer troca de sinalização entre eles. Por exemplo, na Figura 3.1 tem-se $x \text{ co } z$, pois z está fora do alcance do cone de luz de x .

Dois elementos são ditos concorrentes se ambos ocorrem sem influenciar um ao outro. Note que, concorrência não significa necessariamente simultaneidade [19], como pode ser observado na Figura 3.1 tem-se $x \text{ co } z$, portanto nenhuma sinalização pode ser trocada entre eles, e esses eventos não ocorrem ao mesmo tempo. Por outro lado, a relação de causalidade (*li*) entre dois elementos x e y definida como

$$x \text{ li } y \Leftrightarrow x \prec y \text{ ou } y \prec x, \quad (3.3)$$

¹Em um ordenamento total os elementos do conjunto X tem que obedecer, além das três condições de ordenamento parcial, uma quarta condição dada por $\forall x, y$ ou $x \prec y$ ou $y \prec x$.

acontece entre elementos que tem uma dependência causal e portanto suas ocorrências satisfazem a Equação 3.1. Da Figura 3.1 tem-se $x \text{ li } y$.

Exemplos de concorrência ocorrem por exemplo em sistemas relativísticos compostos de corpos em movimento, pois dois corpos podem ter percepções opostas da simultaneidade ou mais geralmente da ordem temporal dos eventos. Outros exemplos são sistemas com múltiplos processadores, sistemas distribuídos e redes de comunicação.

Partindo de axiomas baseados em conceitos da Física, Petri apresentou a Teoria da Concorrência como uma base geral para a teoria de redes² [19]. Ela pode ser concebida como uma combinação de axiomas que incluem causalidade e concorrência. Uma característica essencial da Teoria da Concorrência é que a relação de concorrência co determina univocamente uma rede causal e assim uma ordem causal, que realiza exatamente as restrições de concorrência dadas por co [20].

Uma estrutura de concorrência CS (do inglês, *Concurrency Structure*) é uma tripla que consiste de um conjunto X (que pode ser finito ou infinito) e duas relações binárias, li e co , definidas em X que satisfazem os axiomas da concorrência de Petri [19].

O conjunto X pode ser particionado em dois conjuntos S e T tais que $S \cup T = X$ e $S \cap T = \emptyset$. O conjunto S será relacionado com elementos passivos que são denominados de estados. Esses estados são modificados por elementos do conjunto T que são relacionados com os elementos ativos denominados de eventos. Uma característica das redes é a partição dos seus elementos em dois tipos: lugares que representam elementos passivos e transições que representam elementos ativos. Assim, dada uma estrutura de concorrência $CS = (X, li, co)$, naturalmente faz-se a associação de estados com lugares e de eventos com transições dessa estrutura.

A relação de fluxo no espaço-tempo indica a escolha de uma certa direção no tempo na qual o fluxo (do estado distribuído) é orientado. No nível de concorrência não existe

²Uma rede consiste basicamente de um conjunto de elementos interconectados. Essa interconexão representa alguma relação entre os elementos. Por exemplo dado um conjunto de pessoas é possível formar uma rede usando esse conjunto se dois elementos são conectados se eles têm algum grau de parentesco.

uma direção privilegiada. A partir de uma estrutura de concorrência $CS = (X, li, co)$ é possível encontrar redes com diferentes orientações das relações de fluxo. Analisando a dinâmica dessas redes observa-se que algumas delas são equivalentes as estruturas concorrentes, no sentido que representam o fluxo do estado distribuído. Mostra-se ainda que existe no mínimo uma dessas orientações que é dita ser consistente, que é uma relação F orientando cada par de vizinhos imediatos³ x em y em exatamente uma direção. Do ponto de vista de um elemento x a relação $F^{-1}[x]$ indica os elementos que tem efeito direto em x , e $F[x]$ os elementos que são diretamente afetados por x e $F \cap F^{-1} = \emptyset$.

Além da partição dos seus elementos em lugares passivos e transições ativas, outra característica importante das redes é que elas são objetos matemáticos apropriados para capturar as idéias de estados e mudanças de estados, além da possibilidade de expressar relações de causalidade e concorrência por meio de simples relações de fluxo [21].

Portanto, para modelar uma estrutura de concorrência podem ser usadas as redes. A seguir são definidas um tipo de redes denominadas redes de Petri.

3.2 Redes de Petri

As redes de Petri são uma ferramenta de modelagem matemática com uma representação gráfica [22]. Uma rede de Petri é um grafo direcionado, ponderado e bipartido consistindo de dois tipos de nós, denominados lugares e transições, sendo que arcos podem ser ponderados e são direcionados de lugares para transições ou de transições para lugares. Na representação gráfica, lugares são representados por círculos e as transições por barras ou retângulos, veja Figura 3.2. A cada lugar pode ser associado um inteiro positivo representado graficamente por pontos dentro desse lugar que são denominados de fichas.

³A vizinhança imediata $im[x]$ de algum evento em X contém todos os estados locais que são modificados por esse evento. A vizinhança imediata $im[x]$ de um estado local x contém aqueles eventos que mudam x .

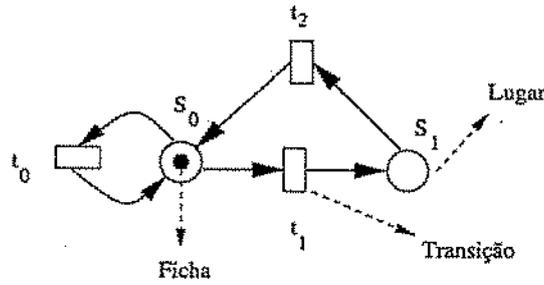


Figura 3.2: Exemplo de rede de Petri.

A seguir tem-se a definição formal das redes de Petri [10].

Definição 3.2 (Redes de Petri Lugar/Transição) *Uma rede de Petri lugar/transição é uma quintupla (P, T, F, W, M_0) , sendo:*

1. $P = \{p_1, p_2, \dots, p_m\}$ é um conjunto finito de lugares,
2. $T = \{t_1, t_2, \dots, t_\ell\}$ é um conjunto finito de transições,
3. $F \subseteq (P \times T) \cup (T \times P)$ é um conjunto de arcos,
4. $W : F \rightarrow \{1, 2, 3, \dots\}$ é a função de ponderação,
5. $M_0 : P \rightarrow \{0, 1, 2, 3, \dots\}$ é a marcação inicial,
6. $P \cap T = \emptyset$ e $P \cup T \neq \emptyset$.

A definição formal de redes de Petri foi feita em termos de redes lugar/transição, que serão algumas algumas vezes simplesmente denominadas de redes de Petri, apesar de existir outros tipos como as redes condição/eventos, redes coloridas, dentre outras. A definição foi feita dessa forma pois este tipo de rede será predominantemente usado nos resultados que serão apresentados mais adiante.

A marcação, denotada por M , é um vetor de m componentes, sendo m o número total de lugares da rede, em particular $M(p_i)$ $i \in \{1, 2, \dots, m\}$ denota que na marcação M o inteiro $M(p_i)$ está associado ao lugar p_i . Nesse caso diz-se que o lugar p_i contém $M(p_i)$ fichas.

Nessa tese, o interesse maior recai sobre o estabelecimento de medidas de complexidade e de capacidade, o que pede a representação explícita de todas as ações concorrentes e de seus efeitos nos sistemas analisados. Por esse motivo, a semântica de passos (*steps*) foi escolhida. Os passos de uma rede são definidos a seguir.

Definição 3.3 (Passos) *O conjunto de Passos⁴ de uma rede de Petri é definido como $S = \{E : E \subseteq T, E \neq \emptyset\}$.*

Deve ser observado que o conjunto vazio de eventos não pode ser considerado um passo válido. Além disso, não se faz qualquer observação a respeito da ocorrência ou não dos passos de uma rede, pois quem definirá suas ocorrências será a regra de disparo definida na próxima seção. Quando os passos consistem somente dos subconjuntos de mesmo tamanho (com a mesma quantidade de eventos, todos habilitados concorrentemente) é dito que neste caso se considera apenas os passos viáveis.

Definição 3.4 (Pré e Pós-conjuntos) *Seja uma rede de Petri (P, T, F, W, M_0) e $x \in X$.*

(1) $\cdot x = \{y \in X \mid (y, x) \in F\}$ é o pré-conjunto de x .

(2) $x^* = \{y \in X \mid (x, y) \in F\}$ é o pós-conjunto de x .

(3) $\cdot x^* = \cdot x \cup x^*$.

A seguir, a regra que rege a ocorrência de um passo denominada de regra de disparo é apresentada, e a mudança de marcação causada pela ocorrência desse passo.

⁴A ocorrência de um passo em uma rede representa a ocorrência de uma ação no sistema modelado por esta rede. O termo evento em redes Petri é usado para representar a ocorrência uma única transição.

3.2.1 Regra de disparo

Dependendo da aplicação, várias interpretações podem ser associadas aos elementos de uma rede de Petri, mas em todos os casos marcações representam estados distribuídos do sistema modelado por essa rede. Neste trabalho, lugares são associados a *estados locais* e as ocorrências das transições a *eventos locais*.

A ocorrência de um evento em um sistema é representada no seu modelo em rede de Petri pelo disparo da transição correspondente aquele evento. De forma análoga, quando todas as transições que formam um passo podem disparar, as condições para isso serão estabelecidas a seguir, diz-se que elas estão habilitadas concorrentemente, o disparo em conjunto dessas transições será denominado de **disparo do passo**. A regra de disparo descrita nessa seção especifica sob que condições uma marcação M habilita concorrentemente um conjunto de transições, e como a ocorrência desse conjunto de transições muda a marcação M para uma nova marcação M' . Os axiomas a seguir [11] definem precisamente as conseqüências da ocorrência de um passo:

A1: Um passo transforma uma marcação M em outra marcação M' .

A2: Passos são a única fonte de mudanças numa rede de Petri.

A3: Um passo E está habilitado se $\forall t_i \in E : p \in {}^*t_i \Rightarrow M(p) \geq W(p, t_i)$.

A4: (Regra de disparo) Se um passo E ocorre em uma marcação M produzindo uma nova marcação M' (representada por $M[E]M'$), então para todo $t_i \in E$, $W(p, t_i)$ fichas são removidas de cada lugar em *t_i e $W(t_i, p)$ fichas são acrescentadas a cada lugar $p \in t_i^*$.

Estes axiomas definem precisamente a noção de *transição em um passo* de uma marcação M para uma nova marcação M' pela ocorrência de um conjunto $E \subseteq T$, isto é, um conjunto de transições *concorrentes* que podem ocorrer numa rede de Petri. Nesse caso diz-se que M' é alcançada a partir de M .

O exemplo a seguir ilustra os conceitos de *transição em um passo* e de *concorrência*.

Exemplo 3.1 A rede de Petri mostrada na Figura 3.3 representa um sistema com dois usuários ou processos (1 e 2) que disputam um mesmo recurso, a disponibilidade do recurso é modelada pela presença de uma ficha no lugar p_4 . O usuário 1 (2) pode acessar o recurso, representado pelo disparo da transição t_2 (t_4), em seguida liberá-lo, disparo de t_3 (t_5). Seguindo a seqüência de disparos de transições do usuário 1, t_2 seguida de t_3 , a marcação apresentada na Figura 3.4 é alcançada. Nesta nova marcação o usuário 1 pode realizar alguma operação que não necessita do recurso, representada pelo disparo de t_1 . Já o usuário 2 pode acessar o recurso, que agora está disponível para ele. Observe que as ocorrências dessas transições representam eventos locais.

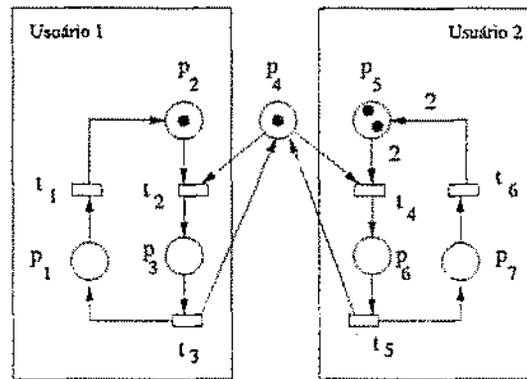


Figura 3.3: Modelo em rede de Petri para o sistema com dois usuários.

Apenas para salientar o efeito da ponderação dos arcos observe que na Figura 3.3 os arcos que saem de p_5 para a transição t_4 , e de t_6 para p_5 têm peso 2, os demais têm peso 1. No caso dos arcos com peso 1 convencionou-se em não escrever esse valor. Ao disparar t_4 duas fichas são retiradas de p_5 e uma é criada em p_6 , analogamente o disparo de t_6 retira uma ficha de p_7 e cria duas em p_5 .

Na marcação apresentada na Figura 3.4, as transições t_1 e t_4 estão habilitadas concorrentemente, portanto essas transições podem ocorrer independentemente ou simultaneamente, assim tem-se os passos $e_1 = \{t_1\}$, $e_2 = \{t_4\}$ e $e_3 = \{t_1, t_4\}$, aqui tem-se um exemplo de **concorrência**. O disparo de qualquer desses passos seguindo a regra

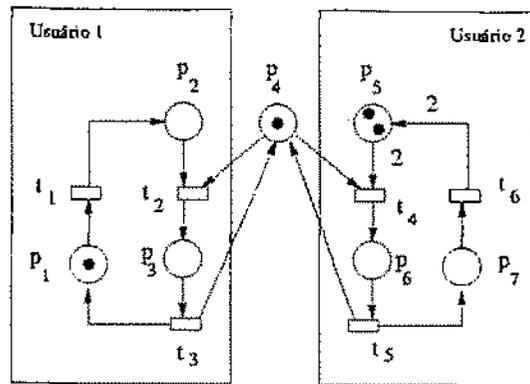


Figura 3.4: Modelo em rede de Petri para o sistema com dois usuários após a seqüência de transições t_2t_3 .

de disparo definida nessa seção muda a marcação da rede e assim tem-se a transição em um passo.

A quantidade de fichas em um determinado lugar da rede é dada por $M(p) = W(t_i, p) - W(p, t_j)$ sendo $i \in {}^*p$ e $j \in p^*$. Portanto, é possível que a quantidade de fichas que é adicionada sempre seja maior que a quantidade retirada, o que implica em um crescimento da quantidade de fichas nesse lugar, se essa quantidade sempre cresce diz-se que a rede é ilimitada. Porém, se em todas as marcações da rede os lugares tem um número máximo de fichas fixo, diz-se que essa rede é limitada.

Na rede da Figura 3.3 é possível observar que as transições t_2 e t_4 podem disparar, nesse caso diz-se que elas estão habilitadas. Ao disparar t_2 a marcação é modificada como ilustrado na Figura 3.5. Nessa nova marcação a transição t_4 não pode mais disparar, pois uma de suas condições de entrada, ficha no lugar p_4 , não é mais satisfeita. Da mesma forma se, partindo da marcação inicial ocorrer o disparo de t_4 , gerando a marcação ilustrada na Figura 3.6, t_2 não estará mais habilitada. Nesse caso, diz-se que essas duas transições estão em **conflito**.

Lembrando que uma marcação M' é alcançável a partir de outra marcação M se existe uma seqüência de passos que após os seus disparos levam de M para M' , se iniciando de M_0 e dos passos que essa marcação habilita forem obtidas as marcações

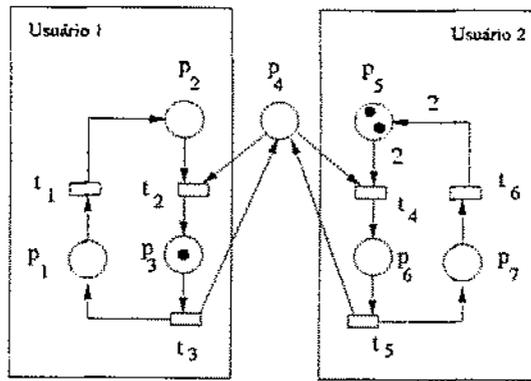


Figura 3.5: Marcação após o disparo de t_2 .

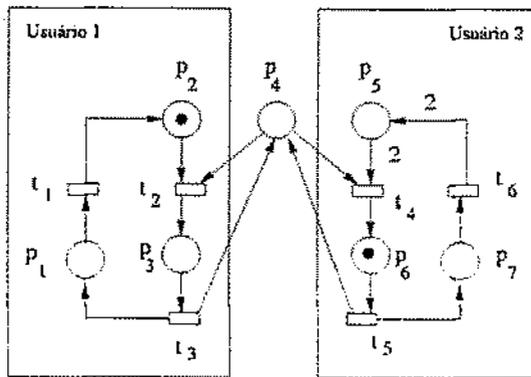


Figura 3.6: Marcação após o disparo de t_4 .

alcançadas, e repetindo esse processo para todas as marcações obtidas tem-se o conjunto de todas as marcações da rede denominado conjunto alcançável, que é denotado por $R(M_0)$ [10, p. 547]. Uma rede de Petri é dita ser viva se independente de qual marcação tenha sido alcançada a partir de M_0 , sempre é possível disparar qualquer transição da rede após uma seqüência de passos.

Para representar a relação entre marcações e passos, seja a seguinte definição.

Definição 3.5 (Grafo de cobertura) O grafo de cobertura $G = (\mathcal{M}, \mathcal{E})$ de uma rede de Petri (P, T, F, W, M_0) é um grafo direcionado (cuja raiz é M_0), e o conjunto de vértices \mathcal{M} é o conjunto de todas as marcações e, $\mathcal{E} \subset S$ o conjunto de arcos conectando dois vértices M, M' tais que $M[E) M'$, sendo que M e M' estão em \mathcal{M} e $E \in \mathcal{E}$.

Para redes limitadas o grafo de cobertura é denominado de grafo de alcançabilidade.

Para construir o grafo de cobertura inicia-se com a marcação M_0 e obtém-se o conjunto das próximas marcações alcançadas diretamente pelos disparo dos passos habilitados em M_0 . Para cada nova marcação, obtém-se os passos habilitados e as marcações alcançadas pelos disparo desses passos. O grafo de alcançabilidade para a rede da Figura 3.3 é ilustrado na Figura 3.7.

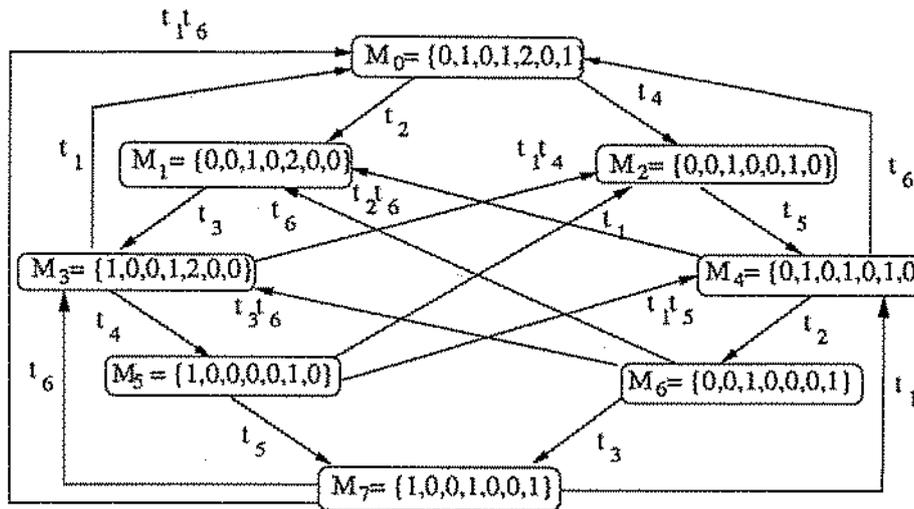


Figura 3.7: Grafo de alcançabilidade para a rede de Petri da Figura 3.3.

Observando a Figura 3.7 é possível observar os passos viáveis $\{t_2, t_6\}$, $\{t_1, t_4\}$, $\{t_1, t_5\}$, $\{t_3, t_6\}$ e $\{t_1, t_6\}$. É possível construir uma grafo de alcançabilidade usando somente os passos viáveis. Esse gafo será usado mais adiante.

Um grafo é determinístico se, em todos os vértices $I \in \mathcal{V}$, os arcos de saída são rotulados distintamente. Dos axiomas A1-A4 tem-se que o grafo de cobertura é determinístico, pois em cada marcação da rede os passos que estão habilitados são distintos e como essas marcações são os vértices dos grafo de cobertura e os arcos são rotulados pelos passos, tem-se que de cada vértice sairá um conjunto de arcos rotulados distintamente.

Se a rede é ilimitada, existem alguns lugares nos quais o número de fichas cresce

indefinidamente, e assim a rede de Petri tem um número infinito de marcações. Uma solução para obter um grafo de cobertura finito, no caso da rede ser ilimitada, é representar um conjunto de marcações cujos vetores, que são marcações da rede, se diferenciam apenas pelo acréscimo de fichas em algumas posições, enquanto outras se mantêm constante. Por exemplo, o conjunto de marcações $M_i = \{1, 0, 1, 0\}$, $M_{i+1} = \{1, 1, 1, 0\}$, $M_{i+2} = \{1, 2, 1, 0\}$, ..., $M_{i+n} = \{1, n, 1, 0\}$, ... podem ser representados pelo vértice $M_i = \{1, \omega, 1, 0\}$, sendo o símbolo ω um inteiro não negativo. Mostra-se em [22] que o número de vértices no grafo de cobertura é sempre finito.

Exemplo 3.2 Uma rede de Petri ilimitada é mostrada na Figura 3.8(a). O grafo de cobertura desta rede é mostrado na Figura 3.8(b).

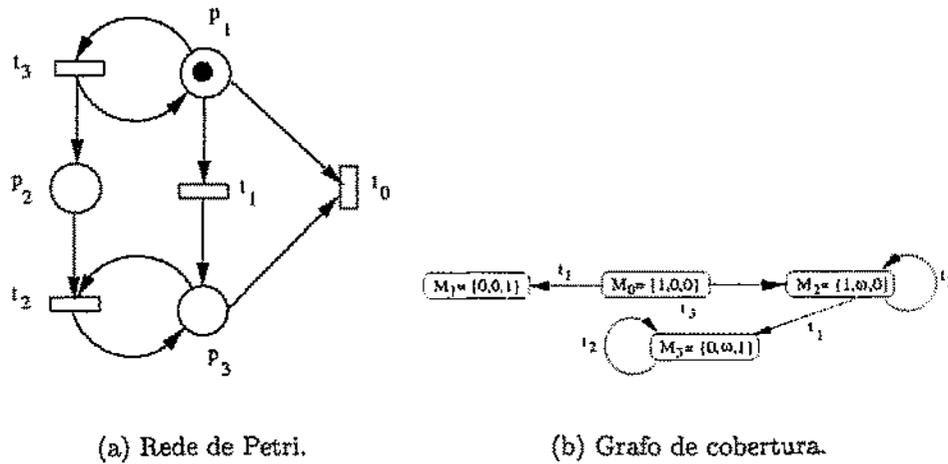


Figura 3.8: Uma rede de Petri ilimitada e seu grafo de cobertura.

Mesmo com a possibilidade da representação compacta do grafo de cobertura nos caso das redes ilimitadas, as rede de Petri estão sujeitas ao problema de explosão de estados [23], pois um sistema com poucos estados locais pode ter muitos ou infinitos estados globais, o que leva a uma grande quantidade de marcações no modelo desse sistema. Existem outras soluções que tentam representar os estados de forma mais

compacta como em Diagramas de decisão binária [24], *Stubborn Sets* [25] e desdobramento da rede (*Unfolding*) [26]. Por não serem usados nesta tese estes métodos não serão descritos.

As redes de Petri também podem ser classificadas como de baixo ou de alto nível. A característica fundamental que classifica uma rede como sendo de alto nível está na possibilidade de associar tipos distintos as fichas, o que não é possível em rede de baixo nível como a rede lugar/transição. Fazendo um paralelo com as linguagens de programação, as rede de baixo nível estão para o *Assembly* assim como as de alto nível estão para as linguagens C ou Pascal. Essa possibilidade de diferenciar o tipo da ficha em um mesmo lugar possibilita às redes de alto nível um poder de compactação dos modelos, como ocorre nas linguagens de programação.

Os resultados desse trabalho foram baseados nas redes lugar/transição, que pertencem a classe de rede de baixo nível, porém foram feitas algumas aplicações usando um tipo de rede de alto nível denominada de rede de Petri Colorida, que será apresentada a seguir.

3.3 Redes de Petri Coloridas

Como extensão das redes de Petri Lugar/Transição, que pertencem a classe de redes denominadas redes de baixo nível, apresentadas na seção anterior, aqui são descritas as redes de alto nível denominadas redes de Petri Coloridas [27]. Vale salientar que os modelos em redes coloridas também podem ser implementados com redes Lugar/Transição, porém em algumas situações o modelo usando esse tipo de rede pode ter muitos lugares e transições como também estruturas repetidas, que podem ser compactadas nas redes Coloridas.

Uma rede de Petri Colorida (CPN, do inglês *Coloured Petri Net*) [27] é uma tupla $\langle \Sigma, P, T, A, N, C, G, E, I \rangle$, que satisfaz os requisitos:

1. Σ é uma conjunto não vazio finito de *cores*,

2. P é um conjunto finito de *lugares*,
3. T é um conjunto finito de *transições*,
4. A é um conjunto finitos de *arcos* tais que $P \cap T = P \cap A = T \cap A = \emptyset$,
5. $N : A \rightarrow P \times T \cup T \times P$ é uma *função de nó*,
6. $C : P \rightarrow \Sigma$ é uma *função cor*,
7. $G : T \rightarrow \text{EXPRESSÕES}$ é uma *função guarda*, tal que

$$\forall t \in T : [\text{Type}(G(t)) = \text{BOOLEAN} \wedge \text{Type}(\text{Var}(G(t))) \subseteq \Sigma],$$

8. $E : A \rightarrow \text{EXPRESSÕES}$ é uma *função de expressão de arco* tal que

$$\forall a \in A : [\text{Type}(E(a)) = C(p(a))_{MS} \wedge \text{Type}(\text{Var}(E(a))) \subseteq \Sigma]$$

sendo $p(a)$ é o conjunto $N(a)$.

9. $I : P \rightarrow \text{EXPRESSÕES}$ é a *função expressão de inicialização* tal que $\forall p \in P :$
 $[\text{Type}(I(p)) = C(p)_{MS} \wedge I(p) \text{ fechado}]$

O conjunto de cores determina os tipos, operações e funções que podem ser usadas nas inscrições da rede (expressões dos arcos, tipos de valores que podem ser colocados em um lugar, dentre outros). Os lugares, transições e arcos tem a mesma definição que nas redes de baixo nível, especificamente nesta tese das redes lugar/transição.

A função nó mapeia cada arco em um par, sendo o primeiro elemento o nó fonte e o segundo o destino e os dois nós de tipos diferentes. A função *cor* mapeia cada lugar p em um conjunto de cor, isto significa que cada ficha no lugar p tem uma cor. A função *guarda* mapeia cada transição em um expressão *booleana*.

As EXPRESSÕES na definição de rede coloridas são as expressões de teste, atribuição ou comparação de valores que podem ser fichas retiradas dos lugares. Deve ser notado

que a sintaxe e a semântica de linguagem usada nas expressões da rede não é apresentada, e é assumida a existência de tal linguagem, de acordo com as mesmas restrições apresentadas em [27] na definição formal de redes de Petri Coloridas. A linguagem CPN-ML como definida em [27] é usada para manipular cores (tipos).

Para o sistema de dois usuários, cujo modelo em rede Lugar/Transição é apresentado na Figura 3.3, considerando que o segundo usuário requer apenas uma ficha para ter acesso ao recurso tem-se o modelo da Figura 3.9. Nos modelos em rede de Petri Colorida, a definição das cores, constantes e as variáveis são feitas no nó de declarações, que na Figura 3.9 está localizado acima do modelo. Tanto o modelo em rede Lugar/Transição como em rede Colorida modelam o comportamento do sistema, porém no segundo caso o modelo é mais compacto.

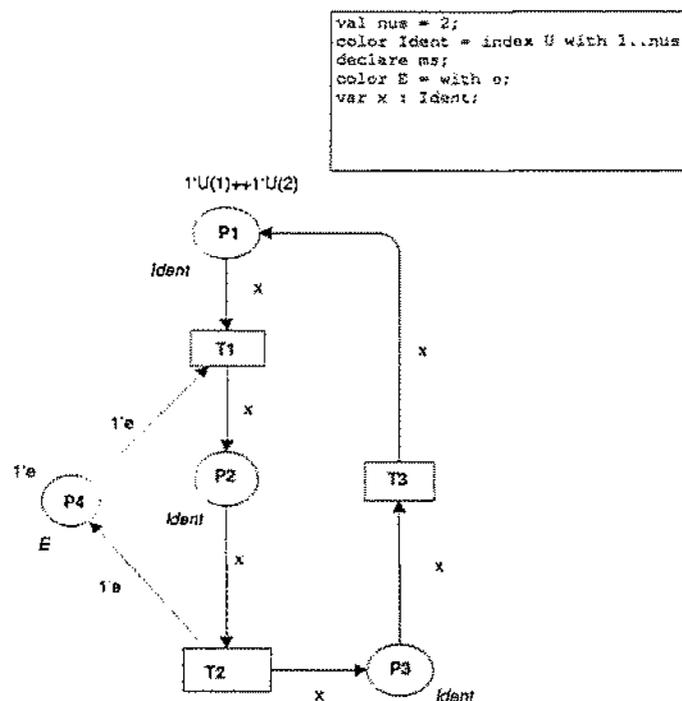


Figura 3.9: Modelo em rede de Petri Colorida para o sistema com dois usuários.

Caso o número de usuários aumente, no modelo em rede Lugar/Transição terão que ser acrescentados os sub-modelos, conjunto de lugares e transições do modelo apresentado

na Figura 3.3 que representam as ações individuais de cada usuário, represente cada usuário. No caso das redes coloridas, por exemplo para o modelo da Figura 3.9, basta modificar a variável nus no nó de declaração e as expressões dos arcos.

Vale salientar que nesta tese estamos interessados em redes de baixo nível portanto não nos aprofundaremos nas redes coloridas. Uma boa referência para aprofundamento nas redes coloridas são os livros de Jensen [27].

Além das redes de Petri existem outras formas de modelar a concorrência, que são apresentadas a seguir.

3.4 Outros Modelos para Sistemas Concorrentes

Qualquer modelo para concorrência está relacionado com o comportamento de um sistema distribuído ao nível de abstração desejado, focando em certos aspectos do comportamento e deliberadamente abstraindo outros [28].

Existem vários modelos para concorrência dentre os quais pode-se citar teoria de traços [29] e estruturas de eventos [30]. Em [28] é possível encontrar outros modelos para concorrência e como eles se relacionam com as redes de Petri. Cada um dos modelos de concorrência modela a sua maneira os relacionamentos causais e concorrentes entre os eventos, para em seguida permitir as análises desejadas.

Um desses modelos parte do princípio que a forma tradicional de descrever o comportamento de um sistema é considerar todas as seqüências de ocorrências de eventos admissíveis, denominadas de traços do sistema. Essencialmente a idéia é obter uma linguagem de um alfabeto cujos elementos são os eventos do sistema.

Mazurkiewicz [29] usou a idéia de traços de um sistema e a ela adicionou uma estrutura de independência entre eventos para obter um modelo para concorrência denominado de teoria dos traços. O comportamento do sistema será descrito em termos de traços, porém como os eventos especificados na estrutura de independência são concorrentes algumas seqüências de eventos são equivalentes. Por exemplo, na rede da Figura 3.10 a partir da marcação mostrada a ocorrência das seqüências $t_2t_6t_3$ e $t_6t_2t_3$

são equivalentes, pois a mesma marcação é alcançada após a ocorrência dessas duas seqüências de transições. Esse fato pode ser observado mais facilmente no grafo de alcançabilidade dessa rede (Figura 3.7), pois a marcação mostrada na Figura 3.10 é a marcação M_4 e após a seqüência de disparos e marcação M_3 é alcançada. Veja que se for dito que a marcação M_3 foi alcançada a partir de M_4 por essa seqüência, não possível dizer a ordem em que as transições t_2 e t_6 ocorreram. Reunindo os traços em classes de equivalência é possível analisar o comportamento do sistema concorrente. Maiores detalhes podem ser encontrados em [29].

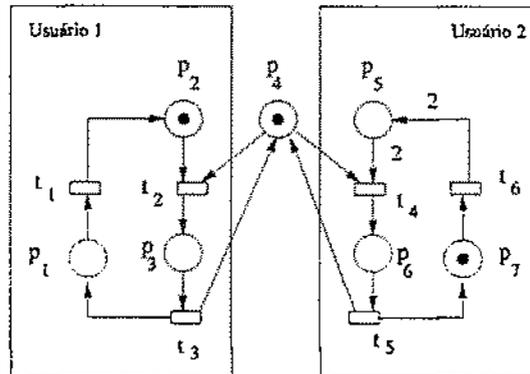


Figura 3.10: Modelo para o sistema com dois usuários e um recurso em comum.

A idéia de traços é similar a teoria de tipos [31]. Nessa teoria, as seqüências de símbolos produzidas por uma fonte ou um sistema são agrupadas de acordo com a freqüência relativa das ocorrências de seus símbolos. Um tipo é um conjunto de todas as seqüências com a mesma freqüência relativa. Mostra-se que observando os tipos é possível calcular a entropia e a capacidade do sistema que gera essa seqüência. No próximo capítulo, será mostrado que a similaridade entre tipos e teoria dos traços foi usada em um dos trabalhos [32] que aplica a teoria da informação em sistemas concorrentes.

Como foi apresentado, além das redes de Petri existem outros modelos para concorrência. Neste trabalho usaremos as redes de Petri. Esta escolha foi guiada pelo tipo de aplicação que é feito, e por resultados prévios relacionados a essas redes e que

serviram de ponto de partida para os trabalhos desenvolvidos que serão apresentados nos próximos capítulos.

3.5 Conclusões

Por fornecerem uma representação gráfica de grande apelo visual e possuir métodos de análise eficientes, as redes de Petri têm se mostrado bastante adequadas ao modelamento de sistemas, principalmente naqueles que apresentam situações que podem ser caracterizadas como conflito ou concorrência.

O desenvolvimento da teoria de redes de Petri, dos métodos para sua análise e o surgimento de outros tipos de rede tais como as redes de Petri coloridas, que permitem a representação compacta de estruturas de dados mais complexas, tem aumentado o poder e a utilidade dessas redes nos mais diversos ramos do conhecimento.

Como será mostrado no próximo capítulo, as rede de Petri serão usadas para o cálculo da capacidade de sistemas que apresentam concorrência. A escolha das redes de Petri em detrimento de outros modelos para concorrência, foi devido a resultados prévios que indicaram ser essas redes a técnica de modelagem mais adequadas para os cálculo da complexidade e capacidade que são os nossos objetivos.

Capítulo 4

Compressibilidade de Redes de Petri e Capacidade de Sistemas Concorrentes

As aplicações da teoria da informação em redes de Petri e *vice-versa* são feitas nesse Capítulo. Inicialmente, usando a idéia de capacidade de canais sem ruído define-se uma medida de complexidade para as redes de Petri. Mostra-se que é possível calcular essa complexidade usando uma matriz de adjacências associada ao grafo de cobertura da rede.

Em seguida, será feita uma aplicação das redes de Petri em teoria da informação. Através do uso dessas redes faz-se o cálculo da capacidade de sistemas com concorrência e com restrição. Um método para o cálculo dessa capacidade será apresentado acompanhado com alguns exemplos.

4.1 Introdução

Segundo Ephremides e Hajek [8] o desenvolvimento dos sistemas de comunicação distribuídos tem sido *ad-hoc*, e isso se deve ao crescimento das redes de comunicação e com isso os problemas que elas apresentam também surgem em maior escala o que leva a maior urgência na sua solução, porém eles enfatizam que técnicas mais elaboradas

para o projeto e análise desses sistemas devem ser desenvolvidas. Até onde vai o conhecimento do autor, no âmbito da teoria da informação, somente recentemente esse desenvolvimento formal foi efetivamente iniciado.

No caso de sistemas concorrentes, Savari [32], [33] usando como base a teoria de traços definiu uma entropia, denominada de entropia de trocas (*interchange entropy*), para as seqüências geradas por uma fonte que produz eventos dentre os quais alguns podem ser concorrentes. O objetivo principal em [32] é analisar um esquema de compressão para as seqüências dessa fonte, para isso as seqüências de símbolos por ela geradas são agrupadas em classes de elementos concorrentes. Mostra-se que no receptor somente é necessário reproduzir uma seqüência que está na mesma classe da seqüência original. Usando essa idéia deseja-se encontrar a taxa de compressão, em bits por símbolo, necessária para representar a classe que contém a seqüência de saída da fonte quando o comprimento da seqüência se aproxima do infinito. É mostrado que a melhor taxa quando o comprimento tende para infinito é a entropia de trocas da fonte.

Os trabalhos de Savari são relacionados com entropia. Aqui o interesse maior é relacionado com complexidade e capacidade. Para tanto, a medida de complexidade para redes de Petri definida por Assis e Barros [34] será estendida. Usando essas redes propõe-se um método para o cálculo da capacidade de um sistema que apresenta concorrência e que tem restrições nas seqüências de símbolos que produz.

4.2 Compressibilidade das Redes de Petri

A seguir, tem-se a definição de uma medida de complexidade para as redes de Petri.

Definição 4.1 *A compressibilidade de uma rede de Petri é dada por*

$$H_{PN} = \lim_{n \rightarrow \infty} \frac{\log_2 N(n)}{n} \quad (\text{bits por passo}) \quad (4.1)$$

sendo $N(n)$ o número de seqüências de comprimento n de passos nessa rede.

O limite em (4.1) existe porque $N(n)$ é uma função sub-aditiva de n [15, p. 104]. A compressibilidade foi inicialmente definida em [34]. Naquele trabalho definiu-se $N(n)$ como a quantidade de seqüências de eventos em redes de Petri vivas e limitadas. Para generalizar essa definição agora são contadas as seqüências de passos, o que expande o significado da compressibilidade pois o caso anterior se torna um sub-caso dessa nova definição. Além disso, não se faz qualquer restrição com relação a vivacidade ou limitação da rede.

Lembrando que numa rede viva a partir de qualquer marcação sempre é possível disparar qualquer transição da rede, e observando que a compressibilidade conta seqüências de passos, tem-se que as transições que não podem ser disparadas não formam passos, e assim não influenciam na contagem da quantidade de seqüências, portanto a compressibilidade também vale para essas redes. Por outro lado, mesmo numa rede não limitada sempre é possível observar as seqüências de passos e contá-las, e dessa forma a compressibilidade também se aplica a essas redes.

É importante notar que o parâmetro H_{PN} é um expoente para a contagem das seqüências de passos de comprimento n no grafo de cobertura, isto é, $N(n)$ cresce aproximadamente com $2^{nH_{PN}}$. Da definição, pode-se dizer que, para um $\epsilon > 0$ arbitrário, existe um inteiro n suficientemente grande tal que

$$2^{n(H_{PN}-\epsilon)} \leq N(n) \leq 2^{n(H_{PN}+\epsilon)}, \quad (4.2)$$

conseqüentemente é possível mapear cada uma das $N(n)$ seqüências de passos em um único bloco de $\lceil n(H_{PN} + \epsilon) \rceil$ bits. Assim, a compressibilidade pode ser interpretada como o número mínimo de bits por passo necessários para representar o grafo de cobertura.

A complexidade descritiva de um objeto foi definida por Kolmokorov [7, p. 144] como o comprimento do menor programa binário de um computador que possa descrever completamente esse objeto. Essa medida de complexidade é importante pois dispensa o uso de probabilidades. Como toda a informação sobre a rede de Petri está contida no seu grafo de cobertura, conclui-se que a compressibilidade também pode ser

interpretada como a complexidade descritiva dessa rede.

Para uma rede de Petri com um conjunto de transições $T = \{t_1, t_2, \dots, t_{|T|}\}$, existem no máximo $2^{|T|} - 1$ passos, pois eles foram definidos como conjuntos não vazios de T . Assim, $N(n) \leq (2^{|T|} - 1)^n$ para todo n , logo:

$$\frac{\log_2 N(n)}{n} \leq \log_2(2^{|T|} - 1), \quad (4.3)$$

então $H_{PN} \leq |T|$. Nos sistemas sem concorrência o conjunto $T = \{t_1, t_2, \dots, t_{|T|}\}$ pode formar no máximo $|T|^n$ seqüências de eventos (não de passos) de comprimento n , então sua compressibilidade é limitada por $H_{PN} \leq \log |T|^n < |T|$.

O valor de $H_{PN} \leq |T|$ indica que para descrever completamente um sistema concorrente é necessário que todas as ações que nele ocorram sejam descritas distintamente. Como nesses sistemas tem-se mais possibilidade de ocorrência de ações, devido a concorrência, do que um sistema com a mesma quantidade de eventos porém sem concorrência, a quantidade de bits necessária é maior.

A observação a seguir é válida para redes de Petri limitadas, isto é, para qualquer lugar p_i , $M(p_i) \leq b$, $b > 0$.

Observação 4.1 Para uma rede de Petri limitada com $|P|$ lugares e conjunto de transições: $T = \{t_1, t_2, \dots, t_{|T|}\}$

$$H_{PN} \leq \min(|T|, |P| \cdot \log_2(b+1)). \quad (4.4)$$

Prova Uma marcação é um vetor $M = (p_1, p_2, \dots, p_{|P|})$, sendo p_i o i -ésimo lugar da rede. Como a rede é limitada tem-se que $0 \leq M(p_i) \leq b$, $i = 1, \dots, m$. Dessa forma, cada lugar pode ser marcado de $b+1$ maneiras. Portanto, se nenhuma restrição for imposta, o número de marcações é igual a

$$(b+1)^{|P|}.$$

Novamente, se nenhuma restrição é imposta às seqüências de passos, um seqüência de n transições pode ocorrer de

$$N(n) = \begin{cases} ((b+1)^{|P|})^n & \text{se } (b+1)^{|P|} \leq (2^{|T|} - 1) \\ (2^{|T|} - 1)^n & \text{se } (b+1)^{|P|} > (2^{|T|} - 1) \end{cases} \quad (4.5)$$

formas. No primeiro caso, existem no máximo $(b+1)^{|P|}$ marcações e como nenhuma restrição foi imposta é possível alcançar de qualquer marcação somente este número de marcações, então somente é possível fazer este número de eventos. No segundo caso em cada marcação é possível disparar $2^{|T|} - 1$ passos, e mais uma vez sem qualquer restrição este será o número máximo de marcações a serem alcançadas e o número máximo de passos que podem ocorrer em cada marcação. Entretanto, da definição de compressibilidade têm-se que

- Se $(b+1)^{|P|} \leq (2^{|T|} - 1)$

$$H_{PN} = \lim_{n \rightarrow \infty} \frac{\log_2 N(n)}{n} \quad (4.6)$$

$$\leq \lim_{n \rightarrow \infty} \frac{\log_2 ((b+1)^{|P|})^n}{n} \quad (4.7)$$

$$= \lim_{n \rightarrow \infty} \frac{n \cdot |P| \cdot \log_2 (b+1)}{n}$$

$$H_{PN} \leq |P| \cdot \log_2 (b+1)$$

- Mas se $(b+1)^{|P|} \geq (2^{|T|} - 1)$ $N(n) = 2^{|T|} - 1$ tem-se o limitante apresentado na observação acima.

■

A seguir é apresentada uma forma de calcular a compressibilidade a partir de seu grafo de cobertura.

4.2.1 Calculando a Compressibilidade

O comportamento assintótico de $N(n)$ para uma rede de Petri pode ser obtida de seu grafo de cobertura. Como o grafo de cobertura tem um número finito de vértices [22],

é possível construir uma matriz finita de adjacências A_{PN} deste grafo (a (i, j) -ésima entrada de A_{PN} é o número de arcos iniciando no vértice i e terminando no vértice j). Usando um argumento similar ao de Shannon [2], a compressibilidade é determinada pelo logaritmo do maior autovalor de A_{PN} .

Exemplo 4.1 *Seja a rede da Figura 3.8 cujo grafo de alcançabilidade está apresentado na Figura 3.7. A matriz de adjacência desse grafo é*

$$A_{PN} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \quad (4.8)$$

cujo maior autovalor é $\lambda = 2,1331$, essa matriz tem outros sete autovalores um é zero e os restantes são complexos, e portanto a compressibilidade do modelo é $H_{PN} = 1,0330$ bits por passo.

É um assunto de investigação futura construir um método de calcular a compressibilidade de uma rede “direto da sua estrutura”. Dessa observação vê-se que as redes de Petri são uma “compactação” bastante eficiente de um sistema concorrente enquanto o seu grafo de cobertura é um tipo de “aproximação de retorno à máquina de estados finitos”.

Arcos no grafo de cobertura de um rede de Petri correspondem a eventos no sistema físico. Porém nem todos os eventos físicos distintos tem informação distinta. Por exemplo, seja um sistema no qual dois usuário a e b podem independentemente (concorrentemente) perfurar um mesmo espaço, veja Figura 4.1. Se a ou b perfura o papel ou eles perfuram em conjunto ab , observando somente a fita vê-se apenas uma perfuração

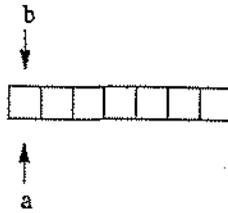


Figura 4.1: Representação de dois usuário escrevendo independentemente em um mesmo papel.

e assim tem-se a mesma informação, portanto, a representação desses eventos deve ser feita atribuindo símbolos idênticos. Para tanto, faz-se necessário atribuir uma função de rotulação ao grafo de cobertura da rede. É também de interesse construir um grafo que caracteriza as seqüências de símbolos sobre um alfabeto \mathcal{A} que obedecem a um conjunto de restrições. Para tratar com ocorrência simultânea de símbolos, tem-se a seguintes definições.

Definição 4.2 (Grafo de cobertura rotulado) *Seja a rede de Petri (P, T, F, W, M_0) com grafo de cobertura $G = (M, \mathcal{E})$. O grafo de cobertura rotulado $G_{\mathcal{A}}$ consiste no par (G, \mathcal{L}) , sendo que a função de rotulação $\mathcal{L} : \mathcal{E} \rightarrow \mathcal{A}$ associa a cada arco (passo) $E \in \mathcal{E}$ um rótulo $\mathcal{L}(E)$ da alfabeto finito \mathcal{A} .*

Observe que o grafo base de $G_{\mathcal{A}}$ é o grafo de cobertura G . Após a rotulação, os caminhos no grafo de cobertura rotulado fornecerão as seqüências que obedecem as restrições modeladas pela rede. O uso do grafo de cobertura para obter as seqüências restritas tem uma diferença fundamental com relação a outros métodos, como por exemplo por diagrama de estados finitos, pois esses grafos contém tanto as seqüências individuais dos eventos que correspondem aos passos com um só elemento como os passos que consistem da ocorrência simultânea de várias transições, que após a rotulação podem receber rótulos distintos dos que foram atribuídos as seus elementos individualmente. Por exemplo, se o passo $e_1 = \{t_1\}$ é rotulado com a e o passo $e_2 = \{t_2\}$ é com b , o passo $e_3 = \{t_1, t_2\}$ pode ser rotulado com c e não necessariamente com ab .

Como a função de rotulação faz a passagem dos eventos físicos para a representação da informação, ela desempenha um papel de grande importância na capacidade de um sistema. Como exemplo, considere um sistema no qual ocorrem n processos concorrentes. Na Figura 4.2 tem-se o modelo em rede de Petri para esse sistema. O grafo de alcançabilidade desse modelo está apresentado na Figura 4.3. Como todos os processos são concorrentes, esse grafo nos mostra que é possível fazer qualquer um dos $2^{|T|} - 1$ passos, e a matriz de adjacências associada a esse grafo tem dimensões 1×1 e é dada por

$$A = [2^{|T|} - 1]$$

que tem maior autovalor $\lambda = 2^{|T|} - 1$ e portanto a compressibilidade da rede é dada por $H_{PN} = \log(2^{|T|} - 1) \approx |T|$.

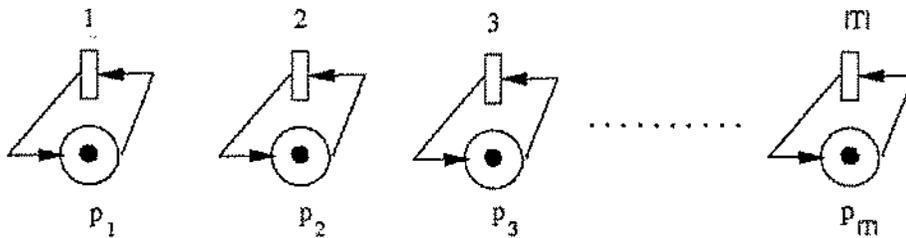


Figura 4.2: Modelo em rede de Petri para $|T|$ eventos concorrentes

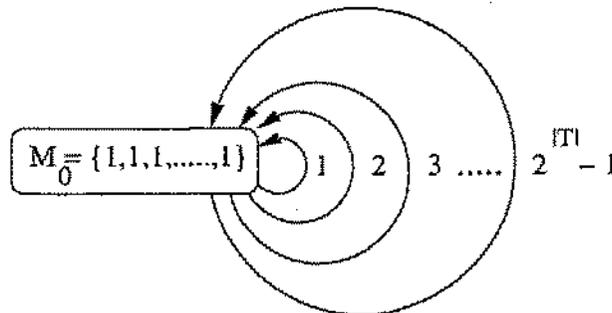


Figura 4.3: Grafo de alcançabilidade para o modelo da Figura 4.2

No sistema da Figura 4.2 já que o ordenamento dos símbolos é impossível devido a concorrência entre esses eventos as seqüências de passos 123, 213, 321, 312, 231 e 132

tem a mesma informação, o mesmo vale para as seqüências de comprimento n formadas pelos mesmo passo. A informação fornecida por essas seqüências pode ser obtida pela observação das quantidades de ocorrência dos símbolos, no exemplo basta dizer que a seqüência tem 1 símbolos de cada elemento do conjunto $\{1, 2, 3\}$.

Uma classificação semelhante é feita na teoria dos tipos [31]. Nessa teoria, as seqüências são estudadas pela observação das freqüências relativas dos símbolos de um alfabeto, por exemplo uma seqüência $x = 123$ cujos símbolos pertencem a um alfabeto $\mathcal{X} = \{1, 2, 3, 4\}$ tem tipo $\mathcal{P}_x = (1/3, 1/3, 1/3, 0)$ e todas as seqüências de tamanho três com essa mesma proporção, por exemplo 321 ou 231 pertencem a esse mesmo tipo. A quantidade de tipos de comprimento l em um alfabeto \mathcal{X} é limitada por

$$|\mathcal{P}_l| \leq (l+1)^{|\mathcal{X}|}. \quad (4.9)$$

Dessa forma, com $|\mathcal{X}|$ eventos é possível formar no máximo $(l+1)^{|\mathcal{X}|}$ passos de comprimento l distintos, e a quantidade de seqüências de comprimento n é dada por $N(n) = |\mathcal{P}_n|$. Substituindo $N(n)$ por $|\mathcal{P}_n|$ na Equação da capacidade tem-se

$$\begin{aligned} C &= \lim_{n \rightarrow \infty} \frac{1}{n} \log N(n) = \lim_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{P}_n| \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \log [(n+1)^{|\mathcal{X}|}] \\ &\leq |\mathcal{X}| \lim_{n \rightarrow \infty} \frac{1}{n} \log (n+1) \\ &= 0. \end{aligned} \quad (4.10)$$

O resultado acima ($C = 0$) para um sistema com todos os eventos totalmente concorrentes entre si é bastante curioso, pois o modelo em rede de Petri desse sistema tem complexidade $H_{PN} = |T|$. Esse fato ocorre pois em um sistema totalmente concorrente não é possível distinguir as seqüências, devido a impossibilidade de ordenação, o que implica em $C = 0$. Porém, para descrever o que ocorre fisicamente é necessário representar explicitamente cada seqüência, o que implica numa complexidade de $H_{PN} = |T|$.

Caso o sistema possa, de alguma forma, apresentar todos os passos do sistema com símbolos distintos, então a capacidade desse sistema será $C = H_{PN} = |T|$. Caso não

seja possível rotular univocamente cada um dos passos, alguns deles receberão rótulos idênticos, o que reduz a capacidade do sistema.

O modelo em rede de Petri de um sistema junto com uma função de rotulação de seu grafo de cobertura, serão usados na próxima seção para obter a capacidade de sistemas concorrentes e com restrições.

4.3 Capacidade de Sistemas Concorrentes com Restrição

Alguns sistemas concorrentes tem que obedecer a restrições nas seqüências que produzem, esses sistemas são definidos a seguir.

Definição 4.3 (Sistemas concorrentes com restrição) *Seja G_A um grafo de cobertura de uma rede de Petri. Um sistema concorrente com restrição sobre o alfabeto A é o conjunto de todas as seqüências obtidas pela leitura dos rótulos de caminhos no grafo G_A .*

A capacidade de um sistema concorrente restringido é definido de forma natural pela Equação

$$C = \lim_{n \rightarrow \infty} \frac{1}{n} \log N(n),$$

sendo $N(n)$ o número de seqüências de subconjuntos permitidos de um alfabeto A .

Como foi apresentado na Seção 2.2.2, o cálculo da capacidade através do maior autovalor da matriz de adjacências do grafo que representa o sistema, requer que esse grafo seja determinístico. É possível que após a rotulação o grafo G_A não seja determinístico, pois mais de um passo pode receber o mesmo símbolo, nesse caso é necessário aplicar um algoritmo que o torna determinístico. Um desses algoritmos está apresentado no Apêndice B.

Propõe-se o seguinte procedimento para o cálculo da capacidade:

1. Construa o modelo em rede de Petri do sistema e obtenha seu grafo de cobertura

2. Defina uma função $\mathcal{L} : \mathcal{E} \rightarrow \mathcal{A}$ que associa a cada passo um dos símbolos de \mathcal{A}

2.1 Se o grafo $G_{\mathcal{A}}$ é determinístico obtenha o maior autovalor da sua matriz de adjacências. A capacidade é o logaritmo desse maior autovalor.

2.2 Se $G_{\mathcal{A}}$ não é determinístico, um algoritmo que produz um grafo determinístico a partir dele, chamado de *construção de subconjuntos* [15, p. 76] apresentado no Apêndice B, pode ser aplicado e a capacidade do sistema é obtida da matriz de adjacência do novo grafo, que agora é determinístico.

O procedimento para a obtenção da capacidade proposto acima depende essencialmente da construção do modelo e da função de rotulação. Até o momento em que esta tese foi finalizada, nenhuma metodologia tinha sido proposta para o modelamento de sistemas de comunicação ou da obtenção da função de rotulação, esse é um assunto para investigação futura.

Na construção do grafo de alcançabilidade os rótulos dos arcos que saem de cada vértice (marcação) são os passos que podem ocorrer naquela marcação. Como os passos em cada marcação são distintos, o grafo de alcançabilidade sempre será determinístico. Se o grafo $G_{\mathcal{A}}$ é determinístico a capacidade de um sistema restringido é igual a compressibilidade do seu modelo em rede de Petri.

A seguir são apresentados exemplos que ilustram os conceitos acima.

Exemplo 1 – Sistema de Transmissão em Múltiplas Trilhas

Bruck e Blaum [35] propuseram um sistema de comunicação assíncrono que consiste de vários canais em paralelo. Nesse sistema, a transmissão em cada canal é feita pela presença ou ausência de um sinal elétrico. A capacidade de um sistema desse tipo com n sub-canais e sem qualquer restrição é $C = n$ bits por uso, pois o transmissor pode enviar em qualquer subconjunto do conjunto de canais. Usando essa idéia como base tem-se o sistema com restrição definido a seguir.

Seja um sistema com vários canais em paralelo. As transmissões em cada canal são feitas por presença (denotado por 1) ou ausência (denotado por 0) de sinais elétricos,

como em [35]. Aqui, tem-se a restrição que após a inserção de uma transição (1) em um canal é necessário ao menos uma ausência de transição (0). Além disso, a ausência de transição simultaneamente no conjunto de canais (0 em ambos os canais) só é permitida ocorrer uma vez. Por exemplo, a seqüência abaixo obedece a restrição desse sistema

```

canal 1  0 0 1 0 0 1 0 1 0 1 0 1 0 0 0 0 1
canal 2  1 0 0 1 0 1 0 0 0 0 0 1 0 1 0 0 1

```

Esse sistema é idêntico a um sistema de gravação em múltiplas trilhas proposto por Marcellin com os parâmetros $d = 1$ e $k = 1$ para duas trilhas. Nesse sistema a restrição d deve ser satisfeita em cada trilha enquanto a restrição k pelo conjunto de trilhas. A capacidade desse sistema é $C = 0,559 \text{ bits/canal}$. Vale observar que o sistema de Marcellin apesar de ser $d = 1$ $k = 1$ no conjunto de trilhas se comporta como $d = 1$ $k = \infty$ em cada trilha. Essa observação tem que ser levada em conta na construção do modelo em rede de Petri, pois o modelo deve ser reduzir à máquina de estados quando se considera apenas um usuário.

O modelo desse sistema em rede de Petri para dois canais é apresentado na Figura 4.4. Nesse modelo as transições t_1 , t_4 , t_2 e t_6 representam ausência de sinal (0), enquanto t_3 e t_5 representam presença (1) de sinal no canal. Veja que após o disparo de t_3 (sinal no canal 1) somente é permitido o disparo de t_2 (ausência de sinal no canal 1), satisfazendo a primeira especificação do sistema. Além disso, se t_3 e t_5 disparam, o que implica no sistema em presença de sinal em ambos os canais, em seguida t_2 , t_6 ou ambos tem que ocorrer, satisfazendo assim a segunda especificação do sistema.

Os passos viáveis desse modelo estão apresentados na Tabela 4.1. O grafo de alcançabilidade dessa rede está apresentado na Figura 4.5 e dele se obtém a Compressibilidade, que nesse caso é igual a capacidade do sistema, igual a $H_{PN} = 1,1178$ que fornece a capacidade por trilha $C = 0,559$ igual ao obtido por Marcellin.

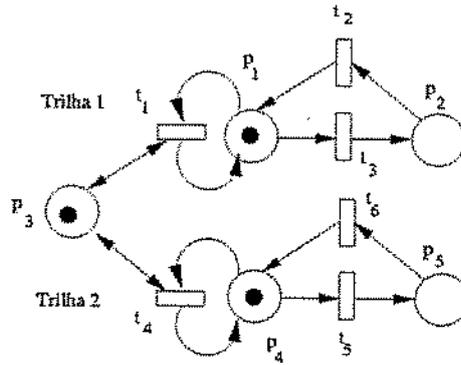


Figura 4.4: Modelo em rede de Petri para o sistema de dois canais em paralelo com restrição.

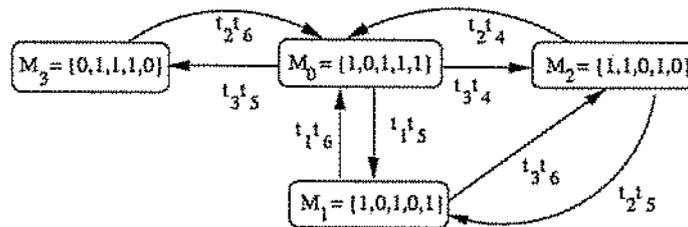


Figura 4.5: Grafo de alcançabilidade do Modelo em rede de Petri para o sistema de dois canais em paralelo com restrição.

Exemplo 2 – Sistema de Comunicação por Cores

Considere um sistema de comunicações no qual o transmissor dispõe de n lâmpadas que emitem luzes de cores distintas. Para enviar as mensagens ele utiliza um subconjunto do conjunto de n luzes. No canal, as luzes são misturadas produzindo uma única cor, e o receptor distingue o subconjunto escolhido pelo transmissor pela cor observada.

Esse sistema, apesar de sua simplicidade, tem as características de um sistema concorrente. No transmissor a ordem em que as luzes foram escolhidas não pode ser completamente descrita, pois independente da ordem que um conjunto de cores seja misturada a mesma cor é obtida. Por ser um sistema concorrente no qual as transmissões ocorrem em paralelo, é possível observar que se nenhuma restrições for imposta, a capacidade de transmissão é de n bits por uso do canal.

Tabela 4.1: Função de rotulação para o grafo de alcançabilidade para os *passos* do modelo da Figura 4.4

<i>Passo</i>	t_1t_5	t_3t_4	t_4t_5	t_1t_6	t_4t_6	t_2, t_3	$\{t_3, t_5\}$
$\mathcal{L}(\textit{Passo})$	(0,1)	(1,0)	(1,1)	(0,0)	(1,0)	(0,0)	(0,1)
<i>Passo</i>	$\{t_2, t_6\}$						
$\mathcal{L}(\textit{Passo})$	(0,0)						

Considere que devido a problemas na detecção das cores no receptor, o transmissor só possa escolher sub-conjuntos de no máximo três cores distintas, veja que agora tem-se um sistema concorrente com restrição. O cálculo da capacidade desse sistema será feito usando redes de Petri.

Para modelar as escolhas entre as n cores sem restrição pode-se usar o modelo apresentado na Figura 4.2 fazendo $|T| = n$ e inserindo um lugar que controle o tamanho máximo do subconjunto do conjunto de cores que podem ser misturadas. O modelo em rede de Petri para $n = 5$ cores com a restrição de que no máximo 3 cores possam ser misturadas é mostrado na Figura 4.6. A quantidade de fichas no lugar p_6 indica o tamanho do maior subconjunto do conjunto de cores que pode ser misturado, no modelo apresentado tem-se que a quantidade máxima é 3. Veja que o passo de tamanho máximo nesse modelo tem três elementos e representa a máxima mistura de cores permitida.

O grafo de alcançabilidade desse modelo é idêntico ao apresentado na Figura 4.3 para $|T| = 5$, porém com menos arcos devido a restrição imposta. Uma função de rotulação para esse sistema é apresentada na Tabela 4.3. A capacidade desse sistema sem qualquer restrição é de $C_0 = \log 32 = 5$.

Após a rotulação do grafo usando a Tabela 4.3 obtêm-se um grafo de ocorrências rotulado determinístico, então a capacidade do sistema é igual a compressibilidade do modelo. Assim tem-se $C_1 = H_{PN} = \log 25 < C_0 = \log 32$.

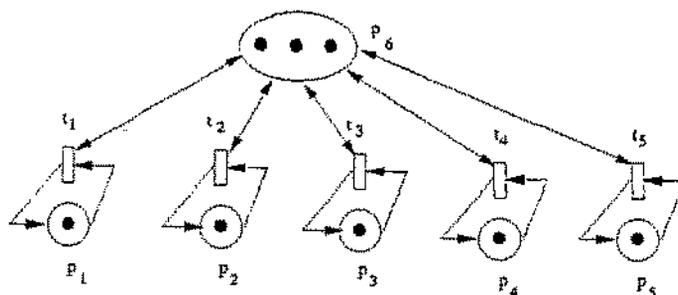


Figura 4.6: Modelo em rede de Petri para o sistema de comunicação por cores.

Considere agora que algumas cores produzidas após as misturas de certos conjuntos de cores não possam ser completamente distinguíveis no receptor. Colocando de outra forma, se todas as misturas que contiverem t_1 e t_3 produzirem a mesma cor então tem-se a rotulação da Tabela 4.3. Nesse caso o grafo rotulado não é mais determinístico e para obter a capacidade tem que ser aplicado o algoritmo de construção por subconjuntos.

Como o grafo de alcançabilidade desse modelo tem apenas um nó, a aplicação do algoritmo de construção por subconjuntos mantém apenas um ramo do conjunto que tem a mesma rotulação. Portanto, para o grafo obtido pela rotulação da Tabela 4.2 dos três arcos com rótulo 6 dois serão retirados e assim a quantidade de vértices será 23 e a capacidade do sistema $C_2 = \log 23 < C_1 < C_0$.

Apesar de sua simplicidade, os exemplos apresentados nessa seção servem para esclarecer os conceitos apresentados, pois em ambos os casos a capacidade pode ser calculada por fórmulas fechadas. A função de rotulação poderia ser aplicada logo na construção do grafo de cobertura usando algoritmos que eliminassem as redundâncias, porém esse é um assunto para investigação futura.

4.3.1 Construção dos Modelos Usando os Passos Viáveis

A construção dos modelos pode ser feita de forma *ad-hoc* ou seguindo algumas regras pré-definidas. No primeiro caso o modelo é construído com base na experiência do projetista, o que não é necessariamente o que ocorre no segundo caso.

Tabela 4.2: Função de rotulação para o grafo de alcançabilidade apresentado na Figura 4.3

<i>Step</i>	t_1	t_2	t_3	t_4	t_5	$\{t_1, t_2\}$
$\mathcal{L}(\text{Step})$	0	1	2	3	4	5
<i>Step</i>	$\{t_1, t_3\}$	$\{t_1, t_4\}$	$\{t_1, t_5\}$	$\{t_2, t_3\}$	$\{t_2, t_4\}$	$\{t_2, t_5\}$
$\mathcal{L}(\text{Step})$	6	7	8	9	10	11
<i>Step</i>	$\{t_3, t_4\}$	$\{t_3, t_5\}$	$\{t_4, t_5\}$	$\{t_1, t_2, t_3\}$	$\{t_1, t_2, t_4\}$	$\{t_1, t_2, t_5\}$
$\mathcal{L}(\text{Step})$	12	13	14	15	16	17
<i>Step</i>	$\{t_1, t_3, t_4\}$	$\{t_1, t_3, t_5\}$	$\{t_1, t_4, t_5\}$	$\{t_2, t_3, t_4\}$	$\{t_2, t_3, t_5\}$	$\{t_2, t_4, t_5\}$
$\mathcal{L}(\text{Step})$	18	19	20	21	22	23
<i>Step</i>	$\{t_3, t_4, t_5\}$					
$\mathcal{L}(\text{Step})$	24					

A seguir tem-se uma primeira formalização de um conjunto de regras para a criação de modelos para os sistemas de comunicação multi-usuários.

- Uma transição representa um e somente um evento local.
- O modelo deve retornar ao sistema não concorrente (se houver) quando se considera apenas um usuário
- Transmissor (fonte) somente podem atuar em transições de uma vizinhança de um conjunto de lugares (em geral um conjunto unitário);
- Receptor (destinatários) somente podem observar as marcações de uma vizinhança de transições;
- Somente passos viáveis são considerados;
- Funções de rotulação podem ser definidas tanto para transições e passos quanto

Tabela 4.3: Outra função de rotulação para o grafo de alcançabilidade apresentado na Figura 4.3

<i>Step</i>	t_1	t_2	t_3	t_4	t_5	$\{t_1, t_2\}$
$\mathcal{L}(\text{Step})$	0	1	2	3	4	5
<i>Step</i>	$\{t_1, t_3\}$	$\{t_1, t_4\}$	$\{t_1, t_5\}$	$\{t_2, t_3\}$	$\{t_2, t_4\}$	$\{t_2, t_5\}$
$\mathcal{L}(\text{Step})$	6	7	8	9	10	11
<i>Step</i>	$\{t_3, t_4\}$	$\{t_3, t_5\}$	$\{t_4, t_5\}$	$\{t_1, t_2, t_3\}$	$\{t_1, t_2, t_4\}$	$\{t_1, t_2, t_5\}$
$\mathcal{L}(\text{Step})$	12	13	14	15	16	17
<i>Step</i>	$\{t_1, t_3, t_4\}$	$\{t_1, t_3, t_5\}$	$\{t_1, t_4, t_5\}$	$\{t_2, t_3, t_4\}$	$\{t_2, t_3, t_5\}$	$\{t_2, t_4, t_5\}$
$\mathcal{L}(\text{Step})$	6	6	20	21	22	23
<i>Step</i>	$\{t_3, t_4, t_5\}$					
$\mathcal{L}(\text{Step})$	24					

para marcações locais observadas pelo destinatário. O nome do símbolo deve ser suficiente para distinção da informação correspondente para o destinatário.

- O sistema sempre retorna à marcação M_0

Exemplo: Canal Acesso Múltiplo Aditivo

Seja o canal com acesso múltiplo ilustrado na Figura 4.7. Nesse canal dois usuários enviam informações x_1 e x_2 sendo $x_1 \in X_1$ e $x_2 \in X_2$, para um receptor usando o mesmo canal. O receptor observando as saídas do canal y sendo $y \in Y$, obterá estimativas \hat{x}_1, \hat{x}_2 das informações transmitidas.

O canal com acesso múltiplo acesso com $X_1 = X_2 = \{0, 1\}$ e $Y = \{0, 1, 2\}$ sendo $y_i = x_1(i) + x_2(i)$ é denominado canal de acesso múltiplo aditivo. A região de capacidade desse canal está apresentado na Figura 4.8.

Seguindo a metodologia proposta na seção anterior obtém-se o modelo da Figura 4.9. Nesse modelo, considera-se que as transições t_1, t_2, t_3 e t_4 representam ações do usuário

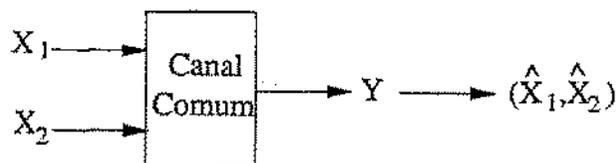


Figura 4.7: Canal com acesso múltiplo.

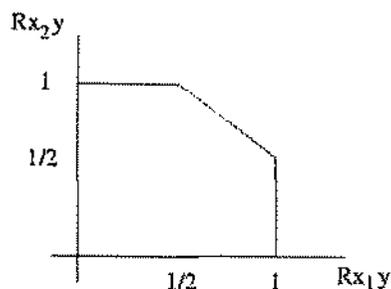


Figura 4.8: Região de capacidade de um canal com acesso múltiplo aditivo.

1 e as transições t_5 a t_8 representam ações do usuário 2. O receptor Y apenas observa as marcações do lugar p_3 . Como exemplo, a ocorrência do passo viável t_1, t_3 representa o envio de zeros pelos dois transmissores. O receptor pode inferir essa informação pela observação da ausência de fichas em p_3 .

Os passos viáveis do modelo apresentado na Figura 4.9 estão listados na Tabela 4.4. O grafo de alcançabilidade do modelo está apresentado na Figura 4.10. Usando esse grafo obtém a Compressibilidade do modelo de 2 bits/passo.

Tabela 4.4: Passos viáveis do modelo apresentado na Figura 4.9.

$E_1 = \{t_1, t_5\}$	$E_2 = \{t_1, t_8\}$	$E_3 = \{t_2, t_5\}$	$E_4 = \{t_2, t_8\}$
$E_5 = \{t_1, t_7\}$	$E_6 = \{t_2, t_7\}$	$E_7 = \{t_1, t_6\}$	$E_8 = \{t_2, t_6\}$
$E_9 = \{t_3, t_5\}$	$E_{10} = \{t_4, t_5\}$	$E_{11} = \{t_4, t_8\}$	$E_{12} = \{t_3, t_8\}$
$E_{13} = \{t_4, t_6\}$	$E_{14} = \{t_4, t_7\}$	$E_{15} = \{t_3, t_7\}$	$S_{16} = \{t_3, t_6\}$

Lembrando que o receptor Y observa apenas a marcação do lugar p_3 que conterà 0, 1 ou 2 fichas. Dessa forma, ele pode inferir que se $M(p_3) = 0$ então $M = M_0$, se

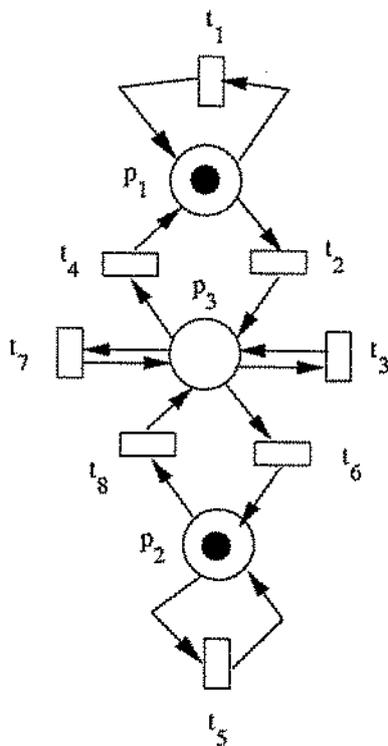


Figura 4.9: Modelo em para o canal MAA.

$M(p_3) = 2$ então $M = M_3$ mas se $M(p_3) = 1$ é impossível decidir se o sistema se encontra em M_1 ou em M_2 .

Então do ponto de vista do receptor Y M_1 e M_2 são indistiguíveis. Nesse caso, o grafo de alcançabilidade pode ser reduzido e assim se obtém o grafo da Figura 4.11. A capacidade obtida desse grafo é $C = \log 3 = 1,5849$.

Considere agora que o segundo usuário fixa a sua transmissão, por exemplo enviando somente zeros, enquanto o primeiro fica livre para transmitir. Aplicando essa informação ao grafo de alcançabilidade da rede deve manter apenas os passos viáveis que refletir essa informação. Fazendo isso obtém-se o grafo da Figura 4.12. A Compressibilidade nesse caso é igual a 1 bit/passo, e esse valor é igual a capacidade quando apenas um dos usuários pode transmitir, conforme pode ser visto na Figura 4.8.

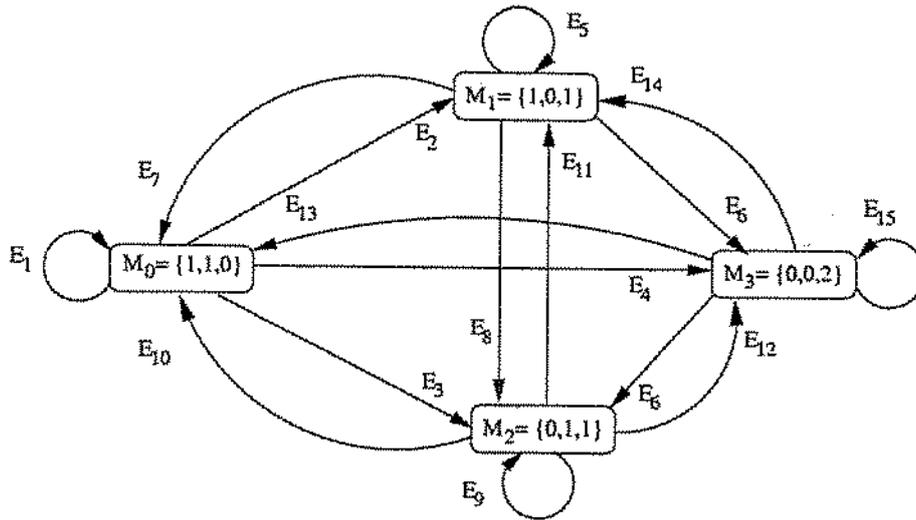


Figura 4.10: Grafo de alcançabilidade para o canal MAA.

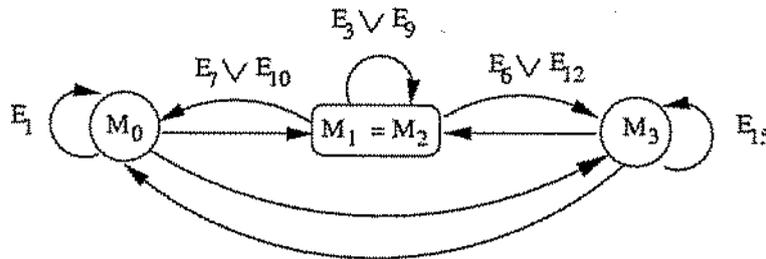


Figura 4.11: Grafo de alcançabilidade reduzido.

4.3.2 Aplicações

Nesta seção, a capacidade de alguns sistemas de gravação é calculada pela compressibilidade do modelo desses sistemas. Mostra-se que o uso do modelo tem vantagens quando comparado com os modelos por diagramas de estados finitos. Para facilitar o modelamento das restrições foram usadas as rede de Petri coloridas.

Capacidade de Sistemas para Gravação em Várias Trilhas

Seqüências discretas que obedecem à restrições na ocorrência de símbolos adjacentes são usualmente empregadas em sistemas de transmissão digital e sistemas de gravação

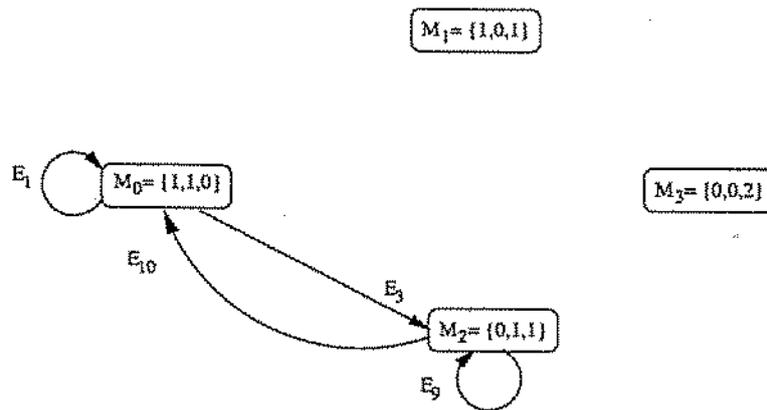


Figura 4.12: Grafo de alcançabilidade reduzido considerando somente as transmissões do primeiro usuário.

óptica e magnética. Entre outras exigências, o uso destas seqüências visa reduzir a interferência intersimbólica e aumentar a habilidade de sincronização de relógio de sistemas de gravação.

As regras que especificam o conjunto de seqüências permitidas são denominadas de *restrições de canal*. O conjunto de todas as seqüências binárias que satisfazem uma restrição de canal será denotado por \mathcal{S} . Por exemplo, o conjunto \mathcal{S} para a restrição de canal unidimensional (d, k) é o conjunto de todas as seqüências binárias tal que o número de 0's entre dois 1's consecutivos é no mínimo d e no máximo k .

Recentes avanços tecnológicos têm possibilitado o aumento da densidade de armazenamento, como por exemplo, o uso de múltiplas cabeças de leitura/gravação operando simultaneamente em n trilhas [36]. Os símbolos lidos/escritos em paralelo em cada trilha são agrupados em um vetor de comprimento n denominado de *vetor de canal*. Um vetor de canal com todas as componentes iguais a zero é denominado de vetor de canal zero. Um DNC bi-dimensional, denotado por DNC-2D, é um canal que admite a transmissão sem ruído de vetores de canal satisfazendo a restrições de ocorrência em intervalos sucessivos.

Marcellin e Weber [37] introduziram um DNC-2D, com parâmetros $(d, k; n)$, onde cada trilha individualmente deve satisfazer a restrição d (número mínimo de 0's entre

dois 1's), enquanto a restrição k limita o número máximo de vetores de canal zero consecutivos. Por exemplo, as seguintes seqüências satisfazem as restrições $(1, 2; 2)$:

```

trilha 1  0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 1
trilha 2  1 0 0 1 0 0 1 0 0 0 0 0 1 0 0 1 0 1

```

No exemplo acima, o primeiro vetor de canal é $[0, 1]$ e o segundo é $[0, 0]$ (vetor de canal zero). Note que embora cada trilha individualmente pode ter seqüências de zeros consecutivos de comprimento maiores do que $k = 2$, o número máximo de vetores de canal zero consecutivos é $k = 2$. Este esquema possibilita o aumento substancial da capacidade de armazenamento em relação ao sistema que lê/grava cada trilha individualmente obedecendo a restrição unidimensional (d, k) em cada trilha [37]. Vários DNC-2D têm sido propostos na literatura [37]-[38].

A capacidade por trilha de um DNC-2D é dada pela Equação 2.5. O método usual para o cálculo da capacidade de um DNC-2D consiste em construir inicialmente o FSD que representa as restrições impostas aos vetores de canal e seguir os procedimentos propostos por Shannon [37]-[38].

Em um dos trabalhos desenvolvidos pelo autor [39], as redes de Petri foram usadas para o cálculo da capacidade de forma sistemática dos códigos $(d, k; n)$ [37] e de uma variante desses códigos denominada de $(c, j; d, k; n)$ [40]. Usando o grafo de ocorrência desses modelos obtém-se a matriz de adjacência e dela a taxa é calculada pelo maior autovalor. Os modelos foram construídos e analisados usando a ferramenta computacional para redes de Petri Coloridas *Design/CPN* [27, 41].

O modelo para o código $(d, k; n)$ é apresentado na Figura 4.13. Esta rede tem dois lugares Trilhas e Contador, representando respectivamente os estados das n trilhas e um contador para controlar a seqüência de vetores de canal zero gravados. A marcação inicial do lugar Trilhas é uma lista ordenada com n fichas do tipo (t, d_t) , $t=1, \dots, n$, sendo t o índice da trilha e d_t a quantidade de símbolos 0's gravados na i -ésima trilha. O último elemento da lista tem de representar o símbolo 1 como o último gravado na trilha n , isto pode ser feito com $(n, 0)$ no fim da lista, pois a marcação

do lugar Contador, que é automaticamente determinada pelos parâmetros k e n foi calculada para esta condição. Por exemplo na Figura 4.13 as marcações podem ser $[(1,0), (2,0)]$ ou $[(1,1), (2,0)]$.

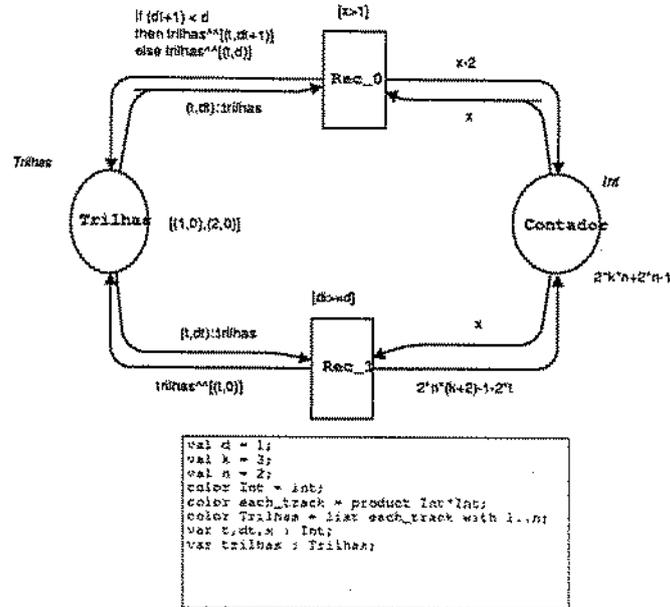


Figura 4.13: Modelo em rede de Petri Colorida para o código $(d, k; n)$.

Existem duas transições Rec_0 e Rec_1 , os disparos dessas transições representam a gravação dos símbolos 0 e 1 respectivamente. Os arcos que conectam o lugar Contador a transição Rec_0 rotulados com x and $x-2$, representam respectivamente as ações de remover todas as fichas de Counter (x), e repor esse número menos 2 fichas ($x-2$) nesse lugar. Da mesma forma, os arcos rotulados com x e $2*n*(k+2)-1-2*t$, que ligam o lugar Contador a Rec_1 , representam respectivamente as ações de remover todas as fichas de Contador (x), e colocar $2*n*(k+2)-1-2*t$ fichas de volta. Assim, para cada disparo de uma transição, a marcação de Contador é atualizada e a ficha (t, d_t) é retirada do início da lista, e a ficha (t, d'_t) com d'_t o novo valor para a quantidade de zeros gravada na trilha t é colocada no final da lista, de tal forma que na próxima ocorrência de uma transição a ficha removida do início da lista representa a próxima trilha a ser gravada.

Como d_t controla o número mínimo de zeros seguidos em cada trilha, este valor somente é atualizado se o valor atual for menor que d , como pode ser observado na expressão do arco de Rec_0 para Trilhas. A transição Rec_1 tem uma guarda [$d_t \geq d$], ela garante no mínimo d 0's após cada 1. A guarda [$x > 1$] na Transição Rec_0 é usada para forçar a ocorrência de Rec_1 quando o número máximo de vetores zero de canal for gravado.

Para obter a taxa dados os parâmetros $(d, k; n)$, basta substituir esses valores no nó de declarações do modelo e construir seu grafo de alcançabilidade. A vantagem em usar este modelo para calcular a capacidade do código, comparado com o procedimento proposto em [37] que basicamente consiste na enumeração de todos os conjuntos de símbolos que podem ser gravados simultaneamente nas n trilhas pra em seguir buscar como esse conjuntos se relacionam de acordo com a restrição, é que aqui o mesmo modelo pode ser usado para todos os valores $(d, k; n)$ sem a necessidade de reconstruir parte do modelo ou usar alguma regra para obter a matriz de adjacência, já que ela é obtida automaticamente.

Para o código $(1, 3; 2)$ obtém-se o grafo de alcançabilidade da Figura 4.14. Essa figura foi obtida usando o *software DESIGN/CPN* que usa para cada vértice a simbologia: índice do vértice seguido da quantidade de arcos que nele incidem, dois pontos a quantidade de arcos que dele partem. Na Tabela 4.5 são listadas as marcações apresentadas no grafo, o rótulo M_i em cada célula da tabela representa o índice i do vértice no grafo e cada marcação tem a lista do lugar Tracks e o valor de counter. Por exemplo, em $M_1 = \{1'[(1, 0), (2, 0)], 1'15\}$ o lugar Tracks tem a lista $[(1, 0), (2, 0)]$ e Tracks tem 15 fichas.

Observando a Figura 4.14 e a Tabela 4.5 vê-se que partindo da marcação M_1 só é possível gravar um zero na trilha 1, isso é feito e em seguida o contador é atualizado e o valor atual $(1, 1)$ é colocado no final da lista, obtendo M_2 . Já em M_3 a trilha 1 pode gravar tanto 0 como 1, restrição $d = 1$ satisfeita, e dessa forma é possível alcançar as marcações M_4 e M_5 .

Considerando com estado final a gravação em todas as n trilhas, que são represen-

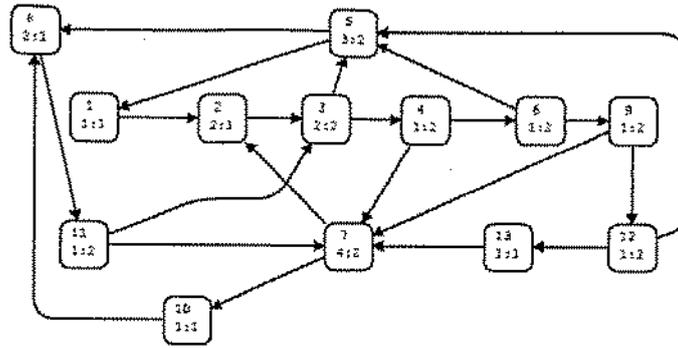


Figura 4.14: Grafo de alcançabilidade para o modelo em rede de Petri para o código (1, 3; 2).

Tabela 4.5: Marcações para o código (1, 3; 2).

$M_1 = \{1'[(1, 0), (2, 0)], 1'15\}$	$M_2 = \{1'[(2, 0), (1, 1)], 1'13\}$
$M_3 = \{1'[(1, 1), (2, 1)], 1'11\}$	$M_4 = \{1'[(2, 1), (1, 1)], 1'9\}$
$M_5 = \{1'[(2, 1), (1, 0)], 1'17\}$	$M_6 = \{1'[(1, 1), (2, 1)], 1'7\}$
$M_7 = \{1'[(1, 1), (2, 0)], 1'15\}$	$M_8 = \{1'[(1, 0), (2, 1)], 1'15\}$
$M_9 = \{1'[(2, 1), (1, 1)], 1'5\}$	$M_{10} = \{1'[(2, 0), (1, 0)], 1'17\}$
$M_{11} = \{1'[(2, 1), (1, 1)], 1'13\}$	$M_{12} = \{1'[(1, 1), (2, 1)], 1'13\}$
$M_{13} = \{1'[(2, 1), (1, 1)], 1'1\}$	

tadas por marcações nas quais a lista está ordenada da primeira para a última trilha, algumas marcações do modelo podem ser consideradas “intermediárias”, por exemplo para o código (1, 3; 2) as marcações M_2 , M_4 , M_5 , M_9 , M_{10} , M_{11} e M_{13} . Aplicando essa regra grafos mais compactos podem ser obtidos. Esses grafos perdem alguma informação a respeito do sistema modelado, porém a informação importante para o cálculo da taxa, que é a taxa de crescimento do número de caminhos é mantida.

Um teste feito para verificar o modelo foi comparar as taxas obtidas com os resultados apresentados em [40]. Foi observado que todos os valores das taxas coincidiram, o que nos deu a segurança de afirmar que, para o nosso objetivo que é usar o modelo

para calcular taxas, o modelo está correto.

A taxa para os códigos $(c, j; d, k; n)$ pode ser obtida usando o modelo para os códigos $(d, k; n)$, removendo do grafo de alcançabilidade as marcações que representam vetores de canal que não obedecem a restrição (c, j) . Por exemplo, para o código $(c = 1, j = 2; d = 1, k = 3; n = 2)$ o grafo ilustrado na Figura 4.14 pode ser usado se forem removidas as marcações que não satisfazem a restrição $c = 1$, obtendo o grafo da Figura 4.15.

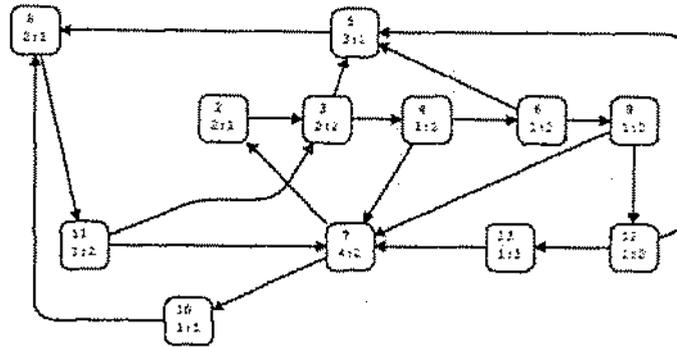


Figura 4.15: Grafo de ocorrência para o modelo em rede de Petri do código $(1, 2; 1, 3; 2)$.

Se a quantidade de vetores de canal é grande, o processo de verificar quais os vetores de canal (marcações) que obedecem a restrição (c, j) pode ser computacionalmente custoso. Usando o modelo da Figura 4.13 é possível obter um modelo para o código $(c, j; d, k, n)$.

Apresentado na Figura 4.16, neste novo modelo dois lugares são adicionados, $Cont_c$ and $Cont_j$, estes lugares são usados, juntamente com as guardas das transições, para implementar as restrições (c, j) em cada vetor de canal. Entretanto, para cada novo vetor de canal as marcações desses novos lugares devem ser re-inicializadas, isto é feito pelo teste no arco de Rec_0 para esses lugares. A marcação inicial de Trilhas tem a mesma forma da que foi usada para o modelo dos códigos $(d, k; n)$, mas deve ser observado que essa lista representa um vetor de canal que satisfaça a restrição (c, j) .

Uma aplicação deste modelo é apresentada na Tabela 4.6. Nesta tabela são apresentados os valores de taxas para alguns códigos. Alguns desses valores não foram cal-

Tabela 4.6: Valores de taxas para alguns códigos $(c, j; d, k; n)$, que não foram calculados em [1].

$(c, j; d, k; n)$	n		
	6	7	8
$(1, n; 2, 7; n)$	0.507	0.507	
$(1, n - 1; 2, 7; n)$	0.429	0.4550	0.4713

4.4 Conclusões

A definição de compressibilidade feita neste capítulo estende uma definição prévia, e aumenta as possibilidades de sua aplicação. Vale observar que resultados obtidos com a definição anterior a apresentada aqui não são invalidados, pois a forma como a compressibilidade era calculada é um sub-caso da definição atual.

A compressibilidade de um modelo representa a quantidade máxima de bits que pode ser usada para representar todas as ações do sistema. Porém, foi mostrado que para obter a capacidade de um sistema concorrente é necessário associar rótulos aos eventos físicos para representar corretamente a informação. Para tanto, foi introduzida uma função de rotulação. Foi mostrado que se essa rotulação é feita de forma que, se todos os eventos do sistema são associadas a símbolos distintos, então a capacidade do sistema é igual a compressibilidade do modelo e é o máximo que se pode fazer em termos de informação.

Capítulo 5

Conclusões

Este trabalho teve como ponto de partida o estudo de sistemas de comunicação multi-usuários, e em especial o cálculo da capacidade nesses sistemas. Os métodos de cálculo da região de capacidade usados atualmente não conseguem determinar exatamente essa região, o que leva a busca de outros métodos para esse cálculo e de análise desses canais. Conclui-se que os eventos nesses sistemas são concorrentes, e para tanto faz-se necessário um método de cálculo da capacidade que considera esse fenômeno.

Para atingir esse objetivo, é desejável obter as medidas de informação apresentadas no Capítulo 2 quando o sistema tem concorrência, para que seja possível analisar qual o efeito da concorrência. Na busca de entender como essas medidas são feitas, e em especial a capacidade, nos sistemas sem concorrência foi desenvolvido o trabalho que está descrito no Apêndice A. Nesse trabalho foram calculados um limitante inferior para a capacidade de um canal com ruído, cujas entradas são seqüências restringidas, e a densidade espectral dessas seqüências. Entretanto, como o foco principal é em sistemas com concorrência, essa linha de trabalho não foi estendida neste trabalho apesar de ser possível estendê-la.

Como os sistemas apresentam concorrência é necessário que se use uma ferramenta que modele a presença desse fenômeno. Algumas ferramentas foram apresentadas e dentre elas as redes de Petri foram escolhidas. Essa escolha se deu por dois motivos

básicos: o primeiro está relacionado a existência de um resultado prévio relacionando uma medida de informação com a complexidade da rede, e o segundo vem da observação que essas redes são adequadas para modelar os sistemas de interesse neste trabalho.

Tendo escolhido o tipo de ferramenta, e usando o conhecimento de resultados já existentes a respeito da aplicação de conceitos de teoria da informação em redes de Petri, obteve-se a primeira contribuição. Originalmente a medida de complexidade estava definida para as redes de Petri vivas e limitadas, e era feita pela contagem de eventos no grafo alcançabilidade dessa rede que fora construído usando o disparo de uma transição por vez. Foi observado que a vivacidade e a limitação da rede não são condições necessárias, e portanto podem ser retiradas. Em seguida, foi observado que a complexidade da rede está diretamente relacionada com a medida de ocorrência dos eventos do sistema que ela modela, e portanto é necessário que se conte todos os eventos, inclusive os que podem ocorrer em paralelo, assim a contagem foi estendida para os passos da rede. Agora, a compressibilidade de um rede de Petri é uma medida de complexidade para qualquer tipo de rede e é feita pela contagem da quantidade de passos da rede.

O cálculo da capacidade de um sistema com restrição pode ser feito usando um modelo que forneça as seqüências de símbolos permitidas. Esse modelo normalmente é feito usando um diagrama de estados finitos. Quando o sistema tem concorrência mostra-se que a representação por esses diagramas é computacionalmente difícil. Nesse ponto tem-se a segunda contribuição desse trabalho.

Usando o modelo em rede de Petri de um sistema cujas seqüências de símbolos tem que obedecer a restrições, foi mostrado que usando uma função de rotulação dos passos da rede nos símbolos de um determinado alfabeto é possível obter a capacidade do sistema. Um método para o cálculo da capacidade foi apresentado juntamente com algumas aplicações. O método aqui proposto junto com algumas aplicações possibilitou a publicação de um artigo no *2003 IEEE International Symposium on Information Theory* [42].

Além das aplicações apresentadas outras foram desenvolvidas como por exemplo

em [43]. Nesse trabalho o cálculo da capacidade de um sistema de transmissão sem fio por irradiação infravermelho definido em [44], cujas seqüências transmitidas tem que obedecer a um conjunto de restrições foi feito usando o método aqui proposto. Usando redes de Petri coloridas foi mostrada a vantagem do uso dessas redes, com relação aos diagramas de estados finitos. Também foi mostrado que no trabalho original [44] que definia o sistema havia um erro no valor da capacidade apresentada.

O método proposto nos forneceu a informação que caso a rotulação consiga representar univocamente todos os passos, a compressibilidade do modelo é igual a capacidade do sistema. Na maioria dos casos a função de rotulação atribui símbolos idênticos a um conjunto de transições, e assim tem-se que a capacidade do sistema diminui e pode-se afirmar que $C \leq H_{PN}$.

O cálculo da compressibilidade é feito pela contagem de caminhos no grafo de cobertura dessa rede. Porém, as redes de Petri sofrem do problema de explosão de estados e assim é possível que o grafo tenha uma quantidade de nós intratável nessa contagem. Uma possibilidade para resolver esse problema consiste em obter um método que calcule a compressibilidade da rede a partir da sua estrutura. Essa é uma das primeiras continuações deste trabalho.

Pode-se questionar em que outros sistemas, além dos sistemas de telecomunicações, o cálculo da capacidade pela compressibilidade seria útil. E assim tem-se mais uma linha de continuação que consiste na busca desses sistemas.

Como foi mostrado durante o texto usamos as redes de Petri como ferramenta para modelar os sistemas concorrentes e que existem outras ferramentas. Será possível usar uma outra ferramenta para os cálculos de capacidade em sistemas com concorrência ?

Apêndice A

Capacidade de Canais M -ários com Restrição (d, k) nas Sequências de Entrada

Neste apêndice é apresentada uma aplicação do cálculo da capacidade de um sistema de comunicações na presença do ruído. Um limitante inferior para a capacidade de um canal com ruído cujas seqüências de entrada são M -árias que obedecem a restrição (d, k) . Mostra-se os efeitos da restrição e do ruído na capacidade do canal.

Uma das motivações do desenvolvimento aqui apresentado foi o de estender resultados encontrados na literatura. Dessa forma, neste apêndice também é calculada a densidade espectral de potências das seqüências M -árias (d, k) .

Vale salientar que os resultados desse apêndice foram apresentados em [45] e que parte dos resultados também foram obtidos independentemente por Datta e MacLaughlin [46].

A.1 Introdução

Nos sistemas de gravação tradicionais, sejam em meios ópticos ou magnéticos, as seqüências binárias devem obedecer a restrições nas quantidades mínima (d) e máxima (k) de zeros entre dois uns, que servem respectivamente para reduzir a interferência inter-simbólica e garantir a recuperação do sincronismo [47]. Os códigos que implementam essas restrições são conhecidos como códigos com restrição nos comprimentos (RLL, do inglês *runlength-limited*) ou mais simplesmente como códigos (d, k) .

Para aumentar a capacidade de gravação, usando o mesmo espaço, novos sistemas que usam seqüências com M -níveis estão sendo desenvolvidos [48]. Como nos materiais usados para a gravação binária, os materiais usados nestes novos sistemas devem manter a linearidade do canal, um requisito necessário para facilitar a recuperação dos dados, e além disso pelos mesmos motivos das seqüências binárias, as seqüências de símbolos M -ários gravadas devem obedecer a restrição (d, k) nos comprimentos. Os códigos que implementam essas restrições são denominados códigos $M(d, k)$. A capacidade de um código $M(d, k)$ para o caso sem ruído é igual ao logaritmo na base-2 da maior raiz real da Equação [2]:

$$x^{k+2} - x^{k+1} - (M - 1)x^{k-d+1} + M - 1 = 0. \quad (\text{A.1})$$

Neste artigo é apresentado um limitante inferior para a capacidade de um canal ruidoso cujas entradas são seqüências M -árias com restrição RLL. O canal ruidoso é modelado como uma generalização de um canal binário simétrico (BSC, do inglês *Binary Symmetric Channel*) que captura o mecanismo de erro que ocorre na recuperação de informações $M(d, k)$ gravadas.

Além da capacidade, outra característica importante dos códigos usados para gravação é a sua densidade espectral de potência (DEP). Dado um código deve ser observado qual o seu nível DC, e se as suas características espectrais podem ser atendidas pelo meio de gravação, o que pode ser feito pela DEP desse código. Neste trabalho a DEP dos códigos $M(d, k)$ é obtida usando a técnica proposta em [49].

A.2 Códigos $M(d, k)$

Um código M -ário com restrição (d, k) (ou simplesmente $M(d, k)$) consiste em um conjunto de seqüências de símbolos de um alfabeto $\mathcal{A} = \{0, 1, 2, \dots, M - 1\}$, nas quais no mínimo d e no máximo k zeros seguidos podem ser aceitos entre os símbolos diferentes de zero. Normalmente, as restrições $M(d, k)$ são representadas por um diagrama de transição de estados finitos (FSTD, do inglês *Finite-State Transition Diagram*) como mostrado na Figura A.1. Qualquer seqüência $M(d, k)$ corresponde a uma seqüência de rótulos obtidos de um caminho nos arcos do FSTD. Existem $(k + 1)$ estados representados pela variável aleatória $S \in \{0, 1, \dots, k\}$.

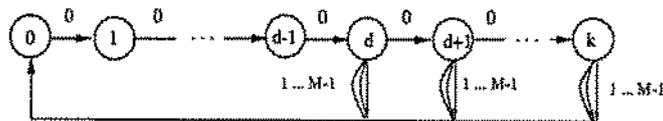


Figura A.1: Diagrama de estados finitos para as seqüências $M(d, k)$.

Denota-se por $P(i, i + 1) = (1 - p_i)$, $i = 0, \dots, k - 1$ as probabilidades de transição do estado i para o estado $i + 1$. Um modelo razoável para a probabilidade de transição do estado i , $i = d, d + 1, \dots, k$ para o estado 0 por um dos arcos que representam a escrita de um dos $M - 1$ símbolos diferentes de zero é

$$P(i, 0) = \frac{p_i}{M - 1}, \quad (\text{A.2})$$

que considera tais símbolos equiprováveis. A probabilidade total de transição do estado i , $i = d, d + 1, \dots, k$ para o estado 0 é portanto p_i . Dessa forma, a matriz de probabilidades de transição correspondente à cadeia de Markov do FSTD mostrado na

Figura A.2 é dada por:

$$\mathbf{P} = \begin{bmatrix}
 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\
 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\
 \vdots & \vdots & \vdots & \dots & 0 & 0 & \dots & 0 \\
 p_d & 0 & 0 & \dots & q_d & 0 & \dots & 0 \\
 p_{d+1} & 0 & 0 & \dots & 0 & q_{d+1} & \dots & 0 \\
 \dots & \dots \\
 p_{k-1} & 0 & 0 & \dots & 0 & 0 & \dots & q_{k-1} \\
 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0
 \end{bmatrix} \quad (\text{A.3})$$

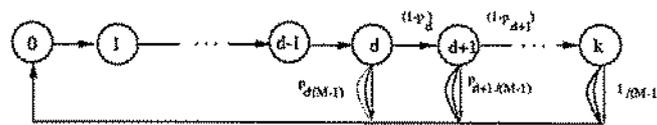


Figura A.2: FSTD para as probabilidades de transição.

A.3 Probabilidades para seqüências M -árias

Em [47] para modelar as seqüências binárias (d, k) foi criada a divisão em frases, em que para uma seqüência de bits o final de cada frase é marcado após o primeiro símbolo 1, e o comprimento da frase é igual a sua quantidade de bits. Foi definida a variável aleatória X_i representando o comprimento das frases e mostrou-se que, sendo C a capacidade calculada pela Equação A.1, a distribuição de probabilidades de X , para a qual a capacidade é alcançada é dada por

$$P(X = i) = 2^{-iC}, \quad i = d + 1, \dots, k + 1. \quad (\text{A.4})$$

Neste trabalho as frases com símbolos $M(d, k)$ são definidas de modo análogo, com o final de cada frase sendo marcado após o primeiro símbolo diferente de zero, como pode ser visto para a seqüência

020300040101023...

cuja divisão em frases é dada por

02, 03, 0004, 01, 01, 02, 3, ...

Definindo as variáveis aleatórias $Y_i = (X_i, A_i)$ em que X_i representa o comprimento da fase i e $A_i \in \{1, 2, \dots, M-1\}$ é o símbolo que marca o final de uma frase, a taxa de entropia (definida no Capítulo 2) é dada por

$$R = \lim_{n \rightarrow \infty} \sup_{P_n(X_1^n, A_1^n)} \frac{H((X_1, A_1), (X_2, A_2), \dots, (X_n, A_n))}{E(X_1 + X_2 + \dots + X_n)} \quad (\text{A.5})$$

em que o supremo é sobre todas as distribuições conjuntas

$$P_n(X_1^n, A_1^n) = P_n((X_1, A_1), \dots, (X_n, A_n)), H((X_1, A_1), (X_2, A_2), \dots, (X_n, A_n))$$

a entropia das primeiras n frases e $E(X_i)$ o valor esperado de X_i . A igualdade em (A.5) é obtida se e somente se as variáveis aleatórias $Y_i = (X_i, A_i)$ e também as variáveis aleatórias A_i forem estatisticamente independentes.

Observação A.1 *As duas condições implicam na independência das variáveis A_i condicionadas às variáveis X_i :*

$$P(A_1, \dots, A_n | X_1, \dots, X_n) = \frac{\prod_{i=1}^n P(X_i, A_i)}{\prod_{i=1}^n P(X_i)} \quad (\text{A.6})$$

$$= \prod_{i=1}^n P(A_i | X_i). \quad (\text{A.7})$$

O resultado a seguir generaliza para o caso M -ário o apresentado em [47, Teorema 1] para sistemas binários.

Proposition 1 *O código que tem taxa máxima de informação tem as seguintes propriedades:*

1. *As variáveis aleatórias X_1, \dots, X_n são estatisticamente independentes e identicamente distribuídas.*

2. As variáveis aleatórias A_1, \dots, A_n são estatisticamente independentes e identicamente distribuídas.

3. A distribuição de probabilidades de X é

$$P(X = i) = (M - 1)2^{-iC}, \quad (\text{A.8})$$

sendo C solução da equação:

$$\sum_{i=d+1}^{k+1} 2^{-iC} = \frac{1}{M - 1}. \quad (\text{A.9})$$

A prova da Proposição 1, segue os passos usados em [47], com algumas modificações.

Prova

Empregando um raciocínio similar ao usado na prova do Teorema 1 em Wolf [47, pag.46] verifica-se que a taxa definida na Equação A.5 é limitada superiormente por:

$$R \leq \sup_{P(X,A)} \frac{H_P(X, A)}{E_{P_X}(X)}, \quad (\text{A.10})$$

em que $P(X, A)$ representa a distribuição da variável aleatória $Y = (X, A)$ e P_X denota a distribuição da variável aleatória X , sendo suprimidos os índices, pois a igualdade em (A.10) é alcançada para variáveis independentes e identicamente distribuídas. Reescrevendo o numerador de (A.10) na forma

$$H_P(X, A) = H_{P_X}(X) + H_{P_{A|X}}(A | X),$$

e considerando a independência de A condicionada à X , vê-se que o máximo valor da entropia $H_{P_{A|X}}(A | X)$ é alcançado para símbolos não nulos equiprováveis, ou seja, $H_{P_{A|X}}(A | X) = \log(M - 1)$, o que está de acordo com o modelo da Figura A.2. Portanto, a taxa de entropia a ser maximizada é

$$R = \frac{H_{P_X}(X) + \log(M - 1)}{E_{P_X}(X)} \quad (\text{A.11})$$

Para calcular P_X^* , a distribuição ótima que maximiza a taxa utiliza-se a técnica dos multiplicadores de Lagrange:

$$F(P_X, \lambda) = \frac{H_{P_X}(X) + \log(M-1)}{E_{P_X}(X)} + \lambda \sum_{i=d+1}^{k+1} P_X(X=i), \quad (\text{A.12})$$

em que a constante λ deve ser obtida a partir da restrição $\sum_{i=d+1}^{k+1} P_X(X=i) = 1$. Denotando $P_i = P_X(X=i)$ e derivando tem-se

$$\begin{aligned} \frac{\partial F(P_X, \lambda)}{\partial P_i} &= \frac{-\left(\log P_i + \log e + i \left(\frac{H_{P_X}(X) + \log(M-1)}{E_{P_X}(X)}\right)\right)}{E_{P_X}(X)} \\ &+ \lambda = 0 \end{aligned} \quad (\text{A.13})$$

A única solução de (A.13) é

$$P^*(X=i) = \frac{2^{-i(C)}}{\sum_{j=d+1}^{k+1} 2^{-j(C)}}, \quad (\text{A.14})$$

em que C é a solução de:

$$\sum_{i=k+1}^{d+1} 2^{-iC} = \frac{1}{M-1}. \quad (\text{A.15})$$

■

É interessante observar que fazendo $2^C = x$, em (A.15), após algumas simplificações obtém-se a Equação A.1, como esperado. Sendo P_X^* a distribuição ótima, dada por (A.14), tem-se

$$R = C = \frac{H_{P_X^*}(X) + \log(M-1)}{E_{P_X^*}(X)}. \quad (\text{A.16})$$

Observe que a capacidade dada pela Equação A.16 leva em consideração a entropia dos comprimentos $H_{P_X^*}(X)$, que independe do último símbolo da frase, uma segunda parcela relacionada com o tamanho alfabeto usado, $\log(M-1)$. Deste modo o comprimento médio das frases é diminuído a medida que M cresce, visto que a capacidade de um sistema M -ário, $M > 2$ deve ser maior que a capacidade de um sistema binário $M = 2$.

Para $M = 2$ a Equação A.15 se reduz para

$$\sum_{i=d+1}^{k+1} 2^{-iC} = 1 \quad (\text{A.17})$$

o que reproduz o resultado apresentado por Wolf [47] para o caso binário.

A ocorrência de erros na leitura dos símbolos $M(d, k)$ é um aspecto importante a ser considerado, um modelo para representar esta ocorrência é apresentado e analisado a seguir.

A.4 Canal com Cruzamento

A Figura A.3 exibe o modelo de canal M -ário sem memória. Denominado de canal com apagamento cruzado (ou simplesmente canal com cruzamento), ele pode representar a recepção ou a leitura de um sinal M -ário cujos símbolos pertencem a um alfabeto \mathcal{A} . Nesse canal um símbolo enviado (ou gravado no nosso caso) pode ser recebido como o seu anterior ou o seu sucessor. Por exemplo, para o alfabeto $\mathcal{A} = \{0, 1, 2, 3, 4\}$ e considerando que símbolo 3 foi enviado, se ocorrer erro é possível detectá-lo como o símbolo 2, ou 4 com uma certa probabilidade.

Seja $Y(S_1, S_2)$ um símbolo de entrada do canal que corresponde ao símbolo M -ário emitido na transição entre S_1 e S_2 , estados consecutivos da máquina de estados que descreve a seqüência de entrada do canal, cada símbolo de saída (Z), com exceção dos símbolos 0 e $M-1$, pode ocorrer devido a presença de três símbolos de entrada, conforme está descrito na Figura A.3.

No modelo a probabilidade de cruzamento é $\alpha/2$, e ele se reduz ao canal binário simétrico para $M = 2$.

Considerando que as seqüências de entrada do canal descrito acima são do tipo $M(d, k)$, na próxima seção será analisado um limitante inferior para a capacidade desse canal.

A.5 Limitante Inferior para a Capacidade

Considere seqüências $M(d, k)$ sendo geradas por uma fonte markoviana e estacionária, e transmitidas através do canal descrito na seção anterior. Nesta seção será calculado

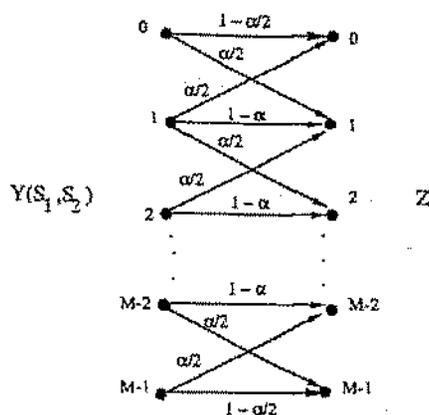


Figura A.3: Canal com apagamento cruzado.

o limitante inferior para a capacidade desse canal sob estas condições.

Em [47] é demonstrado o seguinte lema:

lema 1 *Considere um canal sem memória cuja entrada é um processo aleatório estacionário e markoviano Y , a capacidade deste canal é limitada inferiormente por*

$$C \geq \sup_{P(S_1, S_2)} I(S_2; Z | S_1) \quad (\text{A.18})$$

sendo S_1 e S_2 estados consecutivos do processo markoviano, $P(S_1, S_2)$ a probabilidade de observação dos estados S_1, S_2 consecutivamente e Z a saída do canal perturbada pelo ruído.

A informação mútua da Equação A.18 pode ser escrita como [7]

$$I(S_2; Z | S_1) = H(S_2 | S_1) + H(Z | S_1) - H(S_2; Z | S_1). \quad (\text{A.19})$$

No caso explorado neste trabalho, S_1 e S_2 são estados do FSTD que descreve a seqüência $M(d, k)$ e $Z \in \{0, 1, \dots, M-1\}$. Assim as distribuições de probabilidades para o cálculo das entropias acima podem ser obtidas de análises da Figuras A.2 e A.3. Essas entropias estão apresentadas no Apêndice.

Substituindo as entropias apresentadas no Apêndice na Equação A.19 e após algumas simplificações obtêm-se

$$\begin{aligned}
I(S_2; Z | S_1) = & \\
& \sum_{j=0}^k P[S = j] \{h(p_j) + p_j \log(M - 1)\} \\
& + \sum_{j=d}^{k-1} P[S = j] \left\{ H \left[(1 - p_j) \left(1 - \frac{\alpha}{2}\right) + \frac{p_j}{M-1} \frac{\alpha}{2}, \frac{p_j}{M-1} \left(1 - \frac{\alpha}{2}\right) + (1 - p_j) \frac{\alpha}{2} \right] \right. \\
& \left. - H \left[\left(1 - \frac{\alpha}{2}\right) (1 - p_j), (1 - p_j) \frac{\alpha}{2}, \frac{p_j}{(M-1)} \frac{\alpha}{2}, \left(1 - \frac{\alpha}{2}\right) \frac{p_j}{M-1} \right] \right\}.
\end{aligned} \tag{A.20}$$

Os valores de p_j que representam a probabilidade de retorno de um estado S_j para o estado S_0 são dados por

$$p_j = P(X = j + 1 | X > j) = \frac{P(X = j + 1)}{1 - P(X \leq j)}. \tag{A.21}$$

A probabilidade do estado j é dada por

$$P[S = j] = P[S = 0] \prod_{i=0}^{j-1} q_i, \tag{A.22}$$

sendo $q_j = 1 - p_j$ e $P[S = 0] = 1/L$ em que L é o comprimento médio das seqüências [47], que no nosso caso será calculada usando a distribuição da variável X dada pela Equação A.8.

Observa-se que o limitante obtido para o canal com cruzamento depende do tamanho do alfabeto, o que será explicitado nos limitantes apresentados a seguir.

A.5.1 Exemplos de Limitantes

O cálculo do limitante inferior é realizado usando as Equações A.20 e A.22. O supremo da Equação A.18 é obtido fazendo uma busca no espaço das probabilidades de entrada $p_j, j = d, \dots, k - 1$.

Inicialmente, foi obtido o limitante para $M = 2$, caso binário, para alguns valores de d e k . Conforme o esperado os resultados coincidem com os obtidos em [47].

Para $d = 0$ e $k = 3$, tem-se na Figura A.4 o resultado para alguns valores de M . O mesmo é feito para $d = 1$ e $k = 3$ na Figura A.5.

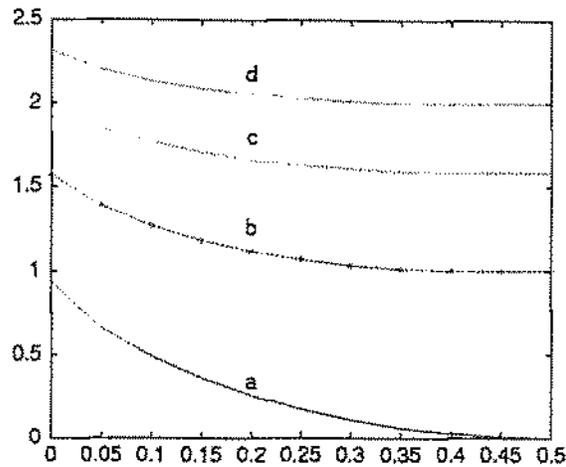


Figura A.4: Limitante inferior da Capacidade em função de α para $d = 0$ e $k = 3$
a) $M = 2$, b) $M = 3$, c) $M = 4$ e d) $M = 5$

Quando $\alpha = 0$, o valor da capacidade deve ser igual ao obtido da Equação A.1 pois não há ruído. Isso foi observado nos resultados obtidos. Também foi observado nas curvas das Figuras A.4 e A.5 o efeito da restrição (d, k) na capacidade, pois quando $\alpha = 0$, a capacidade atinge um valor menor que $C = \log(M)$, que seria a capacidade caso não houvesse a restrição.

Verifica-se um aumento significativo na capacidade do canal M -ário ($M > 2$) sobre o canal binário, ainda que a probabilidade de cruzamento seja elevada.

A.6 Análise Espectral de Cadeias de Markov

Da mesma forma que a capacidade, a densidade espectral de potência (DEP) dos sinais é um dos parâmetros de grande importância nos sistemas de transmissão e gravação. Através da DEP é possível observar se o espectro do sinal a ser gravado é bem ajustado

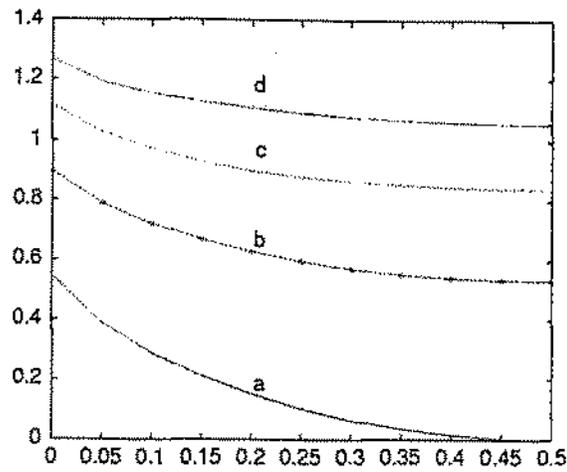


Figura A.5: Limitante inferior da Capacidade em função de α para $d = 1$ e $k = 3$
a) $M = 2$, b) $M = 3$, c) $M = 4$ e d) $M = 5$

ao canal de gravação, o que terá implicações diretas na presença de erros, como também a presença de um nível DC nesse sinal [50].

Para analisar a DEP dos sinais $M(d, k)$, nessa seção é descrito um procedimento para o cálculo da DEP para funções de uma cadeia de Markov. Considerando uma função desse tipo que descreve as seqüências $M(d, k)$ e por conseguinte as seqüências binárias (d, k) fazendo $M = 2$, neste seção é feita a análise espectral dessas seqüências usando o procedimento.

A.6.1 Densidade Espectral de Funções de uma Cadeia de Markov

Para a análise da DEP das seqüências $M(d, k)$, será usado o procedimento desenvolvido em [49]. Nesta seção é feito um resumo deste procedimento.

Considere s_n uma cadeia de Markov com conjunto de estados $S = \{\sigma_1, \sigma_2, \dots, \sigma_I\}$, matriz de probabilidades de transição Π e vetor de probabilidades dos estados \mathbf{p} tal que $\mathbf{p}\Pi = \mathbf{p}$. Seja $a_n = f(s_n)$ um processo estocástico obtido de s_n .

Define-se a matriz

$$\mathbf{A} = \begin{bmatrix} f(\sigma_1) \\ f(\sigma_2) \\ \dots \\ f(\sigma_I) \end{bmatrix}$$

que descreve função $a_n = f(s_n)$. Essa matriz tem dimensões $I \times L$, sendo I a quantidade de estados da cadeia s_n e L a quantidade de componentes do vetor produzido como saída da função $f(\sigma_i)$.

A matriz $\mathbf{\Pi}$ de uma cadeia de Markov ergódica tem as seguintes propriedades:

1. $\mathbf{\Pi}$ tem todos os seus valores não negativos;
2. A soma dos elementos de uma linha é 1

$$\sum_{j=1}^I \Pi_{ij} = 1;$$

3. Essa matriz não tem autovalores com módulo maior que 1, e um desses autovalores é $\lambda_0 = 1$;
4. Se $\mathbf{\Pi}$ tem outros autovalores com módulo 1, $\lambda_1, \dots, \lambda_{N-1}$, esses valores são as raízes da unidade. Se $N > 1$ então $\mathbf{\Pi}$ pode ser escrita como

$$\mathbf{\Pi} = \begin{bmatrix} 0 & \Pi_{12} & 0 & \dots & 0 \\ 0 & 0 & \Pi_{23} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \Pi_{N-1,N} \\ \Pi_{N1} & 0 & 0 & \dots & 0 \end{bmatrix}. \quad (\text{A.23})$$

O valor de N indica a periodicidade da cadeia, que é cíclica se $N > 1$ e acíclica se $N = 1$. Definindo a matriz

$$\mathbf{J}_r = \begin{bmatrix} \mathbf{I}_0 & 0 & \dots & 0 \\ 0 & e^{-j2\pi r/N} \mathbf{I}_1 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & e^{-i2\pi r(N-1)/N} \mathbf{I}_N \end{bmatrix} \quad (\text{A.24})$$

sendo I_0, \dots, I_{N-1} matrizes identidade com as mesmas dimensões das matrizes zero na posição correspondente na Equação A.23. A densidade espectral de potências do processo a_n será dado pela equação

$$S_a(f) = \mathbf{A}^* \mathbf{G}'(e^{-i2\pi f}) \mathbf{C} \mathbf{G}(e^{i2\pi f}) \mathbf{A} \quad (\text{A.25})$$

sendo que \mathbf{G}^* representa transposta e \mathbf{G}' a transposta conjugada da matriz \mathbf{G} e

$$\mathbf{C} = \mathbf{D} - \mathbf{\Pi}' \mathbf{D} \mathbf{\Pi}, \quad (\text{A.26})$$

$$\mathbf{G}(z) = (z\mathbf{I} - \mathbf{H})^{-1}, \quad (\text{A.27})$$

$$\mathbf{D} = \text{diag}\{p(1), p(2), \dots, p(I)\} \quad (\text{A.28})$$

na qual $\text{diag}\{\mathbf{x}\}$ uma matriz cuja diagonal é dada pelos elementos do vetor \mathbf{x} e os demais elementos são zero e $p(i)$ é a probabilidade do estado i .

O valor de \mathbf{H} é dado por

$$\mathbf{H} = \mathbf{\Pi} - \sum_{r=0}^{N-1} e^{i2\pi r/N} \mathbf{J}_r^* \mathbf{w} \mathbf{p} \mathbf{J}_r, \quad (\text{A.29})$$

sendo $\mathbf{w} = [1, 1, \dots, 1]$ um vetor com todos os componentes iguais a 1.

Escrevendo o procedimento do cálculo da DEP em forma de passos, tem-se

1. Escreva a matriz $\mathbf{\Pi}$ no formato da Equação A.23 (somente se s_n é cíclica);
2. Obtenha a matriz \mathbf{A} que descreve a função $a_n = f(s_n)$;
3. Calcule o vetor de probabilidades \mathbf{p} usando $\mathbf{p}\mathbf{\Pi} = \mathbf{p}$;
4. Calcule os valores de \mathbf{H} , \mathbf{G} e \mathbf{C} usando as equações correspondentes;
5. Calcule a DEP usando a Equação A.25.

Caso a cadeia s_n seja acíclica ($N=1$), então $\lambda = 0$ é o único autovalor de Π que está no círculo unitário. Neste caso, a distribuição espectral do processo a_n somente tem linhas espectrais na origem [49]. Se for usado um pulso $h(t)$ para modular a seqüência a_n a DEP do sinal obtido será dada por

$$S(f) = \frac{1}{T} |H(f)|^2 S_a(f). \quad (\text{A.30})$$

O modelo para $a_n = f(s_n)$ no qual o processo a_n é função apenas das transições de estados da cadeia é chamado de máquina de Moore. Outro modelo possível é dado por $a_n = f(s_n, b_n)$ para o qual o processo será função da cadeia e de uma entrada b_n , chamado de máquina de Mealy.

Quando o processo é uma máquina de Moore o cálculo da parte contínua da densidade espectral de potência pode ser feito seguindo o procedimento descrito anteriormente. Porém quando se tem uma máquina de Mealy faz-se necessário algumas modificações no procedimento.

Essas modificações se iniciam introduzindo um novo processo $X_n = (s_n, b_n)$ de tal forma que $a_n = f(X_n)$, e desse modo o problema se transforma em uma máquina de Moore.

Para essa modificação faz-se necessário introduzir algumas variáveis. Sendo $B = \{\beta_1, \dots, \beta_K\}$ o conjunto de entrada com matriz de probabilidades de transição $P = [P(u, v)]$, $u, v = 1, \dots, K$, $A = \{\alpha_1, \dots, \alpha_M\}$ o conjunto de saída, $S = \{\sigma_1, \dots, \sigma_I\}$ o conjunto de estados no qual cada estado é uma possível saída, e g a função de troca de estados que associa o próximo estado em função do atual $S_{n+1} = g(S_n, b_n)$ e definindo a matriz $E_u = [e_u(i, j)]$, $i, j = 1, \dots, I$, $u = 1, \dots, K$ tal que $e_u(i, j) = 1$ se $g(\sigma_i, \beta_u) = \sigma_j$ e $e_u(i, j) = 0$ caso contrário, dois casos podem ser considerados e são apresentadas a seguir.

No primeiro, a seqüência $b_n \in B$ é um conjunto de símbolos independentes e identicamente distribuídos, cujas estatísticas são determinadas por $q_u = P\{b_n = \beta_u\}$

$u = 1, \dots, K$. Neste caso escreve-se a matriz

$$\Pi = \begin{bmatrix} q_1 E_1 & q_2 E_1 & \dots & q_K E_1 \\ q_1 E_2 & q_2 E_2 & \dots & q_K E_2 \\ \vdots & \vdots & \vdots & \vdots \\ q_1 E_K & q_2 E_K & \dots & q_K E_K \end{bmatrix} \quad (\text{A.31})$$

e o vetor A como um conjunto SxB , usando $h(\sigma_1, \beta_u) = x_{(u-1)I+i}$ e aplica-se o método para a máquina de Moore.

No segundo caso, tem-se o processo de entrada b_n como uma cadeia de Markov com matriz de transição $P = P(u, v)$, $u, v = 1, \dots, K$ e definindo $x_{(u-1)I} = (\sigma_i, \beta_u)$, a matriz P será descrita da seguinte forma

$$\Pi = \begin{bmatrix} P(1,1)E_1 & P(1,2)E_1 & \dots & P(1,K)E_1 \\ P(2,1)E_2 & P(2,2)E_2 & \dots & P(2,K)E_2 \\ \vdots & \vdots & \dots & \vdots \\ P(K,1)E_K & P(K,2)E_K & \dots & P(K,K)E_K \end{bmatrix} \quad (\text{A.32})$$

Agora tem-se nos dois casos $a_n = h(X_n)$ e desse modo uma máquina de Moore, portanto sendo possível aplicar o procedimento descrito para esse caso.

A.7 Aplicação aos códigos $M(d, k)$

No caso da análise da densidade espectral de potência das seqüências $M(d, k)$, b_n é a cadeia de Markov que descreve as seqüências de entrada e o conjunto de estados de A serão os M possíveis níveis da seqüência $M(d, k)$. Além dessas considerações faz-se necessário algumas modificações na estrutura da cadeia.

Observando a cadeia de Markov que representa as seqüências $M(d, k)$ ilustrada na Figura A.1, tem-se vários caminhos dos estados S_n , $n = d, \dots, k$ para o estado zero, de tal forma que cada um desses caminhos representa a emissão de um símbolo distinto, e a matriz de transição sendo dada pela Equação A.33.

$$\mathbf{P} = \begin{bmatrix}
 0 & 1 & \dots & 0 & 0 & \dots & 0 \\
 p_d & 0 & \dots & q_d & 0 & \dots & 0 \\
 p_{d+1} & 0 & \dots & 0 & q_{d+1} & \dots & 0 \\
 \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
 p_{k-1} & 0 & \dots & 0 & 0 & \dots & q_{k-1} \\
 p_k & 0 & \dots & 0 & 0 & \dots & 0
 \end{bmatrix} \quad (\text{A.33})$$

Para a aplicação do método faz-se necessário que cada símbolo diferente de zero seja representado unicamente por um caminho de retorno para o estado zero, para que seja possível explicitar os diversos valores gerados através da matriz \mathbf{A} . Essa configuração é obtida aplicando a partição de estados ao estado zero da cadeia, que consiste em particionar o estado s_0 em novos estados $s_{01}, s_{02}, \dots, s_{0i}$ e distribuir os arcos que chegavam e partiam de s_0 nos novos estados [15], obtendo a configuração ilustrada na Figura A.6 cuja matriz de probabilidades de transição é dada pela Equação A.34.

$$\mathbf{P} = \begin{bmatrix}
 0 & \dots & 0 & 1 & \dots & 0 & 0 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 p_d/(M-1) & \dots & p_d/(M-1) & q_d & \dots & 0 & 0 \\
 p_{d+1}/(M-1) & \dots & p_{d+1}/(M-1) & 0 & \dots & 0 & 0 \\
 \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
 p_{k-1}/(M-1) & \dots & p_{k-1}/(M-1) & 0 & \dots & q_{k-1} & 0 \\
 p_k/(M-1) & \dots & p_k/(M-1) & 0 & \dots & 0 & q_k
 \end{bmatrix} \quad (\text{A.34})$$

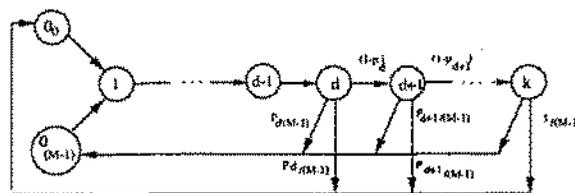


Figura A.6: Aplicação da partição de estados.

A Equação A.34 pode ser usada para modelar o caso em que os símbolos não são identicamente distribuídos, modificando os diversos $p_i/(M-1)$ pelas respectivas

probabilidades. Esse fato não foi analisado nesse trabalho.

Seja a seqüência binária que obedece a restrição $(d, k) = (1, 2)$, representado pela cadeia da Figura A.7, cuja matriz de probabilidades de transição é

$$P = \begin{bmatrix} 0 & 1 & 0 \\ p_1 & 0 & q_1 \\ 1 & 0 & 0 \end{bmatrix} \quad (\text{A.35})$$

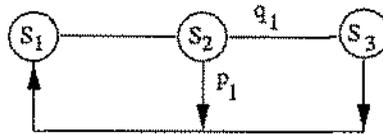


Figura A.7: FSTD para $M = 2$, $(d, k) = (1, 2)$.

Nesse caso, escolhendo os símbolos de saída como -1 e $+1$, faz-se a associação $S = \{\sigma_1 = +1, \sigma_2 = -1\}$ e $B = \{\beta_1 = S_1, \beta_2 = S_2, \beta_3 = S_3\}$.

As matrizes E_i serão $E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ e $E_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, e a matriz de transição Π será

$$\Pi = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ p_1 & 0 & 0 & 0 & q_1 & 0 \\ 0 & p_1 & 0 & 0 & 0 & q_1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{A.36})$$

O vetor de saída será $A = [1 \ -1 \ 1 \ -1 \ 1 \ -1]$.

Para a seqüência acima obtém-se a curva da Figura A.8, que coincide com os resultados gerados em [47].

Agora dois esquemas de codificação para as seqüências $M(d, k)$ serão considerados, e será feita a análise da DEP das seqüências obtidas. O primeiro é uma extensão da codificação NRZI usada em [47], o segundo é um esquema para gravação em meios ópticos proposto em [51].

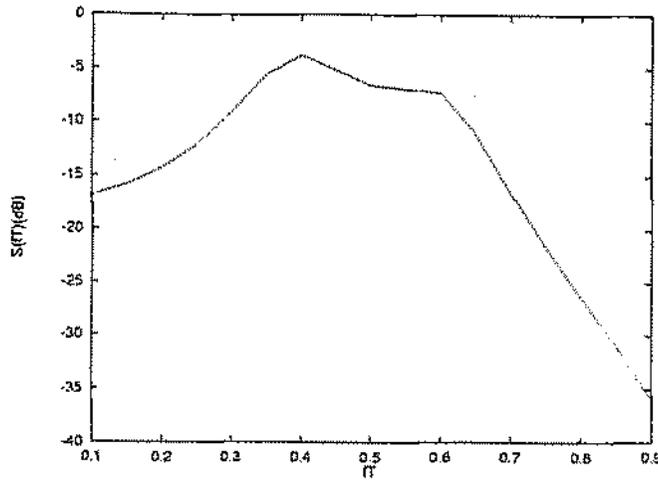


Figura A.8: DEP para $M = 2$ e $(d, k) = (1, 2)$.

Em [47] é proposto um esquema de codificação NRZI a ser aplicado nas seqüências binárias (d, k) . Naquele caso a seqüência resultante pode ter os níveis $+1$ e -1 e, enquanto a entrada for zero a saída permanecerá no nível atual. Como exemplo observe a Figura A.9, nesse caso considera-se que o estado inicial é $+1$ e só haverá troca de nível quando na entrada tivermos o bit 1.

Seqüência (d,k)	0	0	1	0	0	0	1	0	0	1	0	1
NRZI	+1	+1	+1	-1	-1	-1	+1	+1	+1	-1	-1	-1

Figura A.9: Codificação NRZI para o caso binário.

No caso M -ário será feita uma extensão da codificação NRZI. Agora tem-se vários níveis na entrada, e a cada mudança do valor de entrada para um valor $x \neq 0$ a saída terá o valor x , porém o sinal dessa saída será o contrário de uma saída anterior, caso a entrada seja zero ($x = 0$) a saída permanece a mesma, como exemplo observe a Figura A.10.

Na Figura A.11 tem-se a DEP para $(d, k) = (1, 3)$ e alguns valores de M seguindo o esquema NRZI proposto acima. Nota-se que as curvas tem o mesmo formato diferenciando apenas nos valores, pois quanto maior o valor de M , níveis mais altos são

Seqüência $M(d,k)$	0	0	2	0	0	2	0	1	0	1	0	3	
NRZI	+1	+1	+1	-2	-2	-2	+2	+2	-1	-1	+1	+1	-3

Figura A.10: Codificação NRZI para o caso binário.

necessários e mais potência se faz necessário.

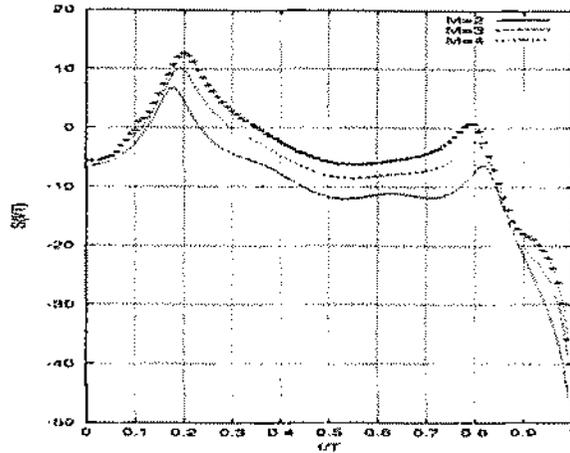


Figura A.11: DEP para $(d, k) = (1, 3)$ e variando o valor de M .

Outro tipo de codificação NRZI é proposto em [51], ele é mais apropriado para meios ópticos, pois a recuperação dos símbolos gravados será feita pela intensidade de um sinal refletido. Nessa caso a seqüência $M(d, k)$ passará por um pré-codificador gerando a saída $y_j = y_{j-1} + x_j \pmod{M}$, como está ilustrado na Figura A.12.

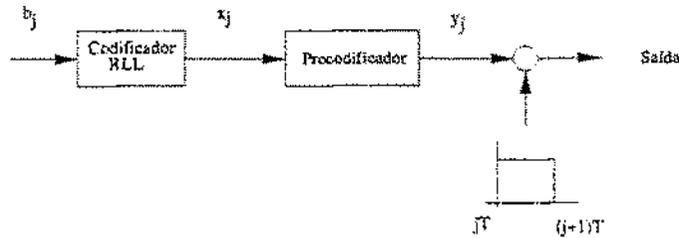


Figura A.12: Esquema NRZI mod M .

Fazendo $(d, k) = (1, 3)$ e para alguns valores de M , tem-se na Figura A.13 a DEP usando o segundo esquema. Nesse caso observa-se a suavização das curvas a medida

que M aumenta.

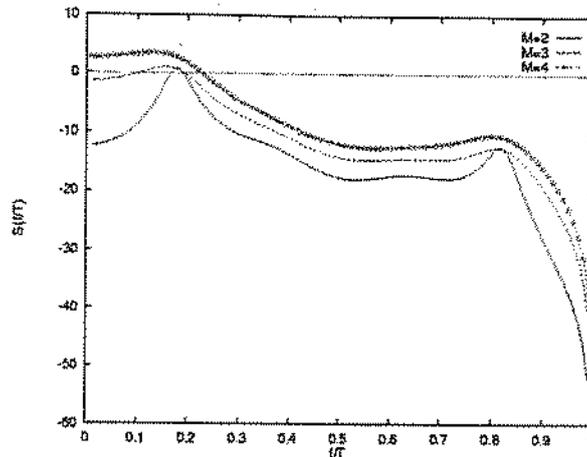


Figura A.13: DEP para o esquema NRZI mod M .

A.8 Conclusões

Neste trabalho os resultados apresentados em [47] para o caso binário foram estendidos para o caso M -ário. Para as seqüência $M(d, k)$ foram obtidas a distribuição de probabilidades para a v.a. que descreve essas seqüências, um limitante inferior para a capacidade de um canal sem ruído cujas seqüências de entrada são desse tipo e por fim a densidade espectral de potência dessas seqüências. Nos três casos fazendo $M = 2$ os resultados obtidos foram comparados com os originais [47] e observou-se a completa coincidência dos valores.

O limitante foi obtido para um canal que é uma extensão do canal BSC. Aqui outro teste foi realizado, pois quando a probabilidade de cruzamento for zero, o canal não tem erros a sua capacidade é igual a da seqüência $M(d, k)$ de entrada que pode ser obtida da Equação A.1, nesse teste observou-se que todos os valores coincidiram. Conforme o esperado, foi observado que o limitante cresce para $M > 2$ e mesmo para níveis de ruído relativamente altos (probabilidade de cruzamento alta) obervou-se um ganho na capacidade a medida que o alfabeto cresce.

No caso da densidade espectral de potência observou-se que ao aumentar o tamanho do alfabeto (M) obtêm-se densidades com formatos idênticos aos obtidos no caso binário para os mesmos d e k . Porém a medida que M cresce essas curvas tem densidades diferentes, embora todas apresentem nível DC baixo.

Em todos os resultados foi considerado símbolos equiprováveis, porém no caso da DEP devido ao método usado existe a possibilidade de testes com alfabetos que não sejam equiprováveis, esses resultados podem ser obtidos em um trabalho futuro.

A.9 Cálculo da Informação Mútua $I(S_2; Z | S_1)$

Neste apêndice são apresentadas as entropias que compõem a informação mútua da Equação A.19, sendo elas $H(S_2 | S_1)$, $H(Z | S_1)$ and, $H(S_2, Z | S_1)$. Denota-se por

$$h(x) = -x \log x - (1 - x) \log(1 - x),$$

a função de entropia binária.

Das Figuras A.2 e A.3 obtêm-se, após algumas análises, as distribuições de probabilidades necessárias para calcular as entropias desejadas, e usando essas distribuições obtêm-se:

$$H(S_2/S_1) = \sum_{j=d}^{k-1} P[S = j] \{h(p_j) + p_j \log(M - 1)\} \quad (\text{A.37})$$

$$\begin{aligned} H(Z/S_1) &= \sum_{j=0}^{d-1} P[S = j] h(\alpha/2) \\ &+ \sum_{j=d}^{k-1} P[S = j] \left\{ H \left[(1 - p_j)(1 - \alpha/2) + \frac{p_j}{M-1} \alpha/2, (1 - p_j)\alpha/2 + \frac{p_j}{M-1} (1 - \alpha/2), \right] \right. \\ &\quad \left. + (M - 2) \left(\frac{p_j}{M-1} \right) \log \left(\frac{p_j}{M-1} \right) \right\} \\ &+ P[S = k] H \left[\frac{1}{M-1} \alpha/2, \frac{1}{M-1} (1 - \alpha/2), \frac{M-2}{M-1} \right]; \end{aligned} \quad (\text{A.38})$$

e

$$\begin{aligned}
H(S_2; Z/S_1) &= \sum_{j=0}^{d-1} P[S = j]h(\alpha/2) \\
&+ \sum_{j=d}^{k-1} P[S = j] \left\{ H \left[(1 - \alpha/2)(1 - p_j), (1 - p_j)\frac{\alpha}{2}, \frac{p_j}{M-1}(1 - \alpha/2), \frac{p_j}{M-1}\alpha/2 \right] \right. \\
&+ (M - 2) \left(\frac{p_j}{M-1} \right) \log \left(\frac{p_j}{M-1} \right) \left. \right\} \\
&+ P[S = k]H \left[\frac{1}{M-1}\alpha/2, \frac{1}{M-1}(1 - \alpha/2), \frac{(M-2)}{M-1} \right]; \tag{A.39}
\end{aligned}$$

Apêndice B

Construção por Subconjuntos (*Subset Construction*)

Neste apêndice é apresentada e exemplificada a construção por subconjuntos que é utilizada para obter um grafo determinístico a partir de outro não determinístico. Essa construção é usada no Capítulo 4.

Dado um grafo rotulado $\mathcal{G} = (G, \mathcal{L})$ que não é determinístico, o objetivo da construção de subconjuntos é encontrar outro grafo $\mathcal{H} = (H, \mathcal{L}')$ que reproduza as mesmas seqüências de símbolos que G e seja determinístico.

A construção inicia fazendo os vértices e I de H como os subconjuntos não vazios, daí o nome construção de subconjuntos, do conjunto de vértices $\mathcal{V}(G)$ de G . Se $I \in \mathcal{V}(H)$ e $a \in \mathcal{A}$ (alfabeto usado para rotular o grafo G), seja \mathcal{J} o conjunto de vértices em G que podem ser alcançados dos elemento de I seguindo os arcos rotulado com a . Em outras palavras, \mathcal{J} é o conjunto de vértices alcançáveis de I usando o rótulo a . Se \mathcal{J} não é vazio, então $\mathcal{J} \in \mathcal{V}(H)$ caso contrário não faça nada. Repetindo o procedimento para cada $I \in \mathcal{V}(H)$ e para todo $a \in \mathcal{A}$ tem-se o grafo rotulado \mathcal{H} . Como para cada vértice $I \in \mathcal{H}$ os arcos que saem tem rótulos distintos a apresentação \mathcal{H} é right-resolving e o o grafo rotulado \mathcal{H} é determinístico.

Example 1 *Seja o a grafo rotulado não determinístico \mathcal{G} da Figura B.1(a). Após a*

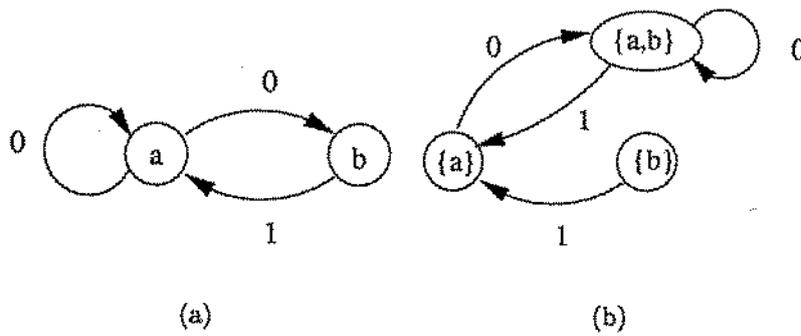


Figura B.1: Aplicação da construção por subconjuntos.

aplicação da construção de subconjuntos obtém-se o grafo rotulado determinístico \mathcal{H} da Figura B.1(b).

B.0.1 Observações

Observe que o grafo da Figura B.1(b) é redutível e para o nosso caso, o cálculo da capacidade, basta usar a parte irredutível do grafo. Portanto, para o cálculo da capacidade a matriz de adjacências deve ser construída observando apenas as adjacências entre os vértices da parte irredutível do grafo obtido após a aplicação da construção de subconjuntos, no caso do exemplo deve-se usar apenas os nós $\{a\}$ e $\{a, b\}$.

O novo grafo rotulado \mathcal{H} , obtido após a construção por subconjuntos a partir de um grafo \mathcal{G} , terá no mínimo a mesma quantidade de vértices de \mathcal{G} pois caso esse último grafo seja determinístico a aplicação da construção de subconjuntos fornecerá um grafo \mathcal{H} com os mesmo vértices de \mathcal{G} . Isso ocorre porque nesse caso os subconjuntos serão apenas os vértices \mathcal{G} e não haverá nenhum vértice de \mathcal{H} que seja combinação dos vértices de \mathcal{G} .

Bibliografia

- [1] Gregory L. Silvus and B.V.K. Kumar. "A Simple Two-Dimensional Run-Length-Limited Code for Ultra-High-Density Magnetic Recording". *Proceedings of International Communications Conference - ICC'95*, pages 633–637, 1995.
- [2] Claude E. Shannon. *A Mathematical Theory of Communication in Key Papers on Information Theory*. IEEE Press, New York, 1961.
- [3] Sergio Verdú. "Fifty Years of Shannon Theory". *IEEE Transactions on Information Theory*, 44(6):2057–2078, October 1998.
- [4] Claude E. Shannon. *Two-way Communication Channels in Key Papers on Information Theory*. IEEE Press, New York, 1961.
- [5] Edward C. Van Der Meulen. "A Survey of Multi-Way Channels in Information Theory: 1961-1976". *IEEE Transactions on Information Theory*, IT-23(1):1–37, January 1977.
- [6] H. B. Meeuwissen. *Information Theoretical Aspects of Two-way Communication*. Technische Universiteit Eindhoven, Eindhoven, 1998.
- [7] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., New York, USA, 1991.
- [8] Anthony Ephremides and Bruce Hajek. "Information Theory and Communication

- Networks: An Unconsummated Union". *IEEE Transactions on Information Theory*, 44(6):2416–24345, October 1998.
- [9] Carl Adam Petri and E. Smith. "Concurrency and Continuity". *Advances in Petri Nets 1987, LNCS 266*, pages 273–292, September 1988.
- [10] Tadao Murata. "Petri Nets: Properties, Analysis and Applications". *Proceedings of the IEEE*, pages 541–580, April 1989.
- [11] C. A. Petri. Nets, time and space. *Theoretical Computer Science*, (153):3–48, 1996.
- [12] Francisco M. de Assis e Marcelo S. Alencar. "Limitante Inferior para a Capacidade de Canais M-ários com Restrição de Comprimento nas Sequências de Entrada". *Anais do Simpósio Brasileiro de Telecomunicações*, pages 155–157, Setembro 1997.
- [13] Robert Ash. *Information Theory*. John Wiley & Sons, Inc., New York, 1967.
- [14] Jack Keil Wolf. "Magnetic Recording as a Communication Channel". *Proceedings of the IEEE International Symposium on Information Theory*, page 5, June 1994.
- [15] Douglas Lind and Brian Marcus. "An Introduction to Symbolic Dynamics and Coding". Cambridge University Press, New York, USA, 1995.
- [16] Nagabhushana T. Sindhushayana. *Symbolic Dynamics, Automata Theory and The Theory of Coding: A Comparative Study and Applications*. Faculty of Graduate School of Cornell University, Master Thesis, 1992.
- [17] J. Pieter Schalkwijk. "The Binary Multiplying Channel - A Coding Scheme that Operates beyond Shannon's Inner Bound Region". *IEEE Transactions on Information Theory*, IT-28(1):107–110, January 1981.
- [18] Carl Adam Petri. "Concurrency Theory". *Petri Nets: Central Models and Their Properties, Lecture Notes on Computer Science 1986*, pages 4–23, September 1987.

- [19] Mark-Oliver Stehr. *“Concurrency Theory of Cyclic and Acyclic Process”*. Universität Hamburg, 1996.
- [20] Olaf Kummer and Mark-Oliver Stehr. “Petri’s Axioms of Concurrency A Selection of Recent Results”. *“Proceedings of the 18th International Conference on Application and Theory of Petri, Lecture Notes in Computer Science 1248, Springer-Verlag”*, pages 1–20, June 1997.
- [21] P.S. Thiagarajan. “Elementary Net Systems”. *Petri Nets: Central Models and Their Properties, Lecture Notes in Computer Science 254, Springer-Verlag*, pages 26–59, September 1986.
- [22] Wolfgang Reisig. *Petri Nets. An Introduction*. Springer-Verlag, Berlin, Alemanha, 1985.
- [23] A. Valmari. “The State Space Explosion Problem”. *“Lecture Notes on Petri Nets I: Basic Models, Springer Verlag”*, pages 429–528, January 1998.
- [24] R. E. Bryant. “Graph-Based Algorithms for Boolean Function Manipulation”. *IEEE Trans, Computers*, C-58(8):677–691, October 1986.
- [25] A. Valmari. “Stubborn sets for reduced state space generation”. *LNCS 489 Advances on Petri Nets 1990*, pages 491–515, 1991.
- [26] Kenneth L. McMillan. *“Symbolic Model Checking”*. Kluwer Academic Published, Boston, 1993.
- [27] Kurt Jensen. *Coloured Petri Nets : Basic Concepts, Analysis Methods and Practical Use (Monographs in Theoretical Computer Science a Series of Eactcs)*. Springer Verlag, 1992.
- [28] Mogens Nielsen and Vladimiro Sassoen. “Petri Nets and Other Models of Concurrency”. *Lecture Notes on Petri Nets I: Basic Models. Advances on Petri Nets*, pages 587–641, 1998.

- [29] A. Mazurkiewicz. "Trace Theory". *Petri Nets: Central Models and Their Properties, Lecture Notes in Computer Science 254 - PART II, Springer-Verlag*, pages 279–324, 1987.
- [30] G. Winskel. "Event Structures". *Petri Nets: Central Models and Their Properties, Lecture Notes in Computer Science 254 - PART II, Springer-Verlag*, pages 325–392, 1987.
- [31] I. Csizsár and J. Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, New York, USA, 1981.
- [32] Serap A. Savari. "On compressing interchange classes of events in a concurrent system". In *Data Compression Conference, 2003. Proceedings. DCC 2003*, pages 153–162, July 2003.
- [33] Serap A. Savari. "Concurrent Process and Interchange Entropy". In *2003 IEEE International Symposium on Information Theory. ISIT 2003*, page To appear, July 2003.
- [34] Francisco M. de Assis and Tomaz C. Barros. "Evaluation of Petri Nets Compressibility". *1997 IEEE International Symposium on Information Theory*, page 323, June 29 - July 4 1997.
- [35] Mario Blaum and Jehoshua Bruck. "Coding for Skew-Tolerant Parallel Asynchronous Communications". *IEEE Transactions on Information Theory*, 39(2):379–388, March 1993.
- [36] L. C. Barbosa. "Simultaneous detection of readback signals from interfering magnetic recording tracks using array heads". *Transactions on Magnetics*, 26(26):2163–2165, September 1990.
- [37] M. W. Marcellin and H. J. Weber. "Two-dimensional modulation codes". *IEEE Journal on Selected Areas in Communications*, 10(3):254–266, September 1992.

- [38] A. Kato and K. Zeger. "On the Capacity of Two-Dimensional Run-Length Constrained Channels". *IEEE Transactions on Information Theory*, 45(45):1527–1540, July 1999.
- [39] Edmar C. Gurjão, Francisco M. de Assis, Angelo Perkusich, Tomaz C. Barros e Cecilio Pimentel. Cálculo da Capacidade de Códigos com Restrições Bi-dimensionais Usando Redes de Petri. *Simpósio Brasileiro de Telecomunicações*, pages 1–5, September 2001.
- [40] Robert A. Swanson and Jack K. Wolf. "A New-Class of Two-Dimensional RLL Recording Codes". *IEEE Transactions on Magnetics*, 28(6):3407–3416, November 1992.
- [41] Kurt Jensen and et all. *Design/CPN 4.0*. Meta Software Corporation and Department of Computer Science, Univeristy of Aarhus, Denmark. On-line Version: <http://www.daimi.aau.dk/designCPN/>, 1999.
- [42] Edmar C. Gurjão, F. M. de Assis, A. Pekusich and Cecilio Pimentel. "Petri Nets Compressibility and Capacity of Concurrent Systems". In *2003 IEEE International Symposium on Information Theory. ISIT 2003*, page To appear, July 2003.
- [43] Edmar C. Gurjão, Francisco M. de Assis and Angelo Perkusich. Some Comments on Petri Net Compressibility with an Application to Calculation of Infrared Wireless Channel Capacity. *I Brazilian Petri Nets Meeting. Evento realizado junto ao XIV Congresso Brasileiro de Automática*, pages 1–5, September 2002.
- [44] Walter Hirt, Martin Hassner, and Nyles Heise. "IrDA-VFIR (16 M/bs): Modulation Code and System Design". *IEEE Personal Communications*, 44(6):58–71, February 2001.
- [45] Edmar C. Gurjão e Francisco M. de Assis. On the Capacity of Noisy Channel with M -ary Runlength-Limited Inputs. *Anais do Simpósio Brasileiro de Telecomunicações*, pages 1–5, Setembro 2001.

- [46] Suprana Datta and Steven W. McLaughlin. Optimal Block Codes for M -ary Runlength-Constrained Channels. *IEEE Transactions on Information Theory*, 47(5):2069–2078, July 2001.
- [47] Ephraim Zehavi and Jack K. Wolf. “On Runlength Codes”. *IEEE Transactions on Information Theory*, 34(1):45–54, January 1988.
- [48] Steven W. McLaughlin, Jian Luo and Quo Xie. On the Capacity of M -ary Runlength-Limited Codes. *IEEE Transactions on Information Theory*, 41(5):1508–1511, September 1995.
- [49] G. L. Pierobon G. Bilardi, and R. Padovani. “Spectral Analysis of Functions of Markov Chain with Applications”. *IEEE Transactions on Communications*, 7(31):853–860, July 1983.
- [50] John G. Proakis. *Digital Communications*. McGraw-Hill Book Company, New York, 1989.
- [51] Steven W. McLaughlin. “Five Runlength-Limited Codes for M -ary Recording Channels”. *IEEE Transactions on Magnetics*, 33(3):244–2452, 1997.