

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM DE SISTEMAS COMPLEXOS

LUIS AMATO NETO

Coincidências: Problema dos aniversários e polinômios cromáticos

São Paulo

2020

LUIS AMATO NETO

Coincidências: Problema dos aniversários e polinômios cromáticos

Versão corrigida

Dissertação apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Modelagem de Sistemas Complexos.

Área de concentração:
Sistemas Complexos

Versão corrigida contendo as alterações solicitadas pela comissão julgadora em 10 de dezembro de 2019. A versão original encontra-se em acervo reservado na Biblioteca da EACH-USP e na Biblioteca Digital de Teses e Dissertações da USP (BDTD), de acordo com a Resolução CoPGr 6018, de 13 de outubro de 2011.

Orientador:
Prof. Dr. José Ricardo Gonçalves de Mendonça

São Paulo

2020

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

CATALOGAÇÃO-NA-PUBLICAÇÃO

(Universidade de São Paulo. Escola de Artes, Ciências e Humanidades. Biblioteca)

CRB 8 - 4936

Amato Neto, Luis

Coincidências: problema dos aniversários e polinômios cromáticos / Luis Amato Neto ; orientador, José Ricardo Gonçalves de Mendonça – 2020.

76 f : il.

Dissertação (Mestrado em Ciências) - Programa de Pós-Graduação em Modelagem de Sistemas Complexos, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, em 2019.

Versão corrigida.

1. Redes complexas. 2. Sistemas discretos. 3. Polinômios.
I. Mendonça, José Ricardo Gonçalves de, orient. II. Título.

CDD 22.ed.– 530.131

Dissertação de autoria de Luís Amato Neto, sob o título ”**Coincidências: Problema dos aniversários e polinômios cromáticos**”, apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Modelagem de Sistemas Complexos, na área de concentração Sistemas Complexos, aprovada em 10 de dezembro de 2019 pela comissão julgadora constituída pelos doutores:

Prof. Dr. José Ricardo Gonçalves de Mendonça

Instituição: Escola de Artes, Ciências e Humanidades da Universidade de São Paulo
Presidente

Prof. Dr. Raphael Yokoingawa de Camargo

Instituição: Centro de Matemática, Computação e Cognição da Universidade Federal do ABC

Prof. Dr. Alexandre Freire

Instituição: Escola de Artes, Ciências e Humanidades da Universidade de São Paulo

Prof. Dr. Camilo Rodrigues Neto

Instituição: Escola de Artes, Ciências e Humanidades da Universidade de São Paulo

AGRADECIMENTOS

Agradeço aos meus pais pela satisfação de aprender em todos os momentos.

À minha esposa Léa que me incentiva e compartilha comigo uma vida de crescimento mútuo.

Às minhas filhas Mariana e Júlia que são alegrias na minha vida.

Ao meu orientador Prof. Dr. José Ricardo por compartilhar um desafio e ter sabedoria e paciência de acompanhar uma longa jornada.

Ao meu colega e amigo Fábio Reale pelas conversas e apoio nos estudos.

Aos professores de Modelagem de Sistemas Complexos pelo empenho em construir um ambiente de aprendizado aberto e acolhedor.

A todos os colaboradores e alunos da EACH que fazem com que tenhamos orgulho de participar do desafio de construir uma Escola.

"Sometimes, a bit of quantitative thinking can show that things are not so surprising after all."

(Parsi Diaconis)

RESUMO

AMATO NETO, Luís. **Coincidências**: Problema dos aniversários e polinômios cromáticos. 2020. 76 f. Dissertação (Mestrado em Ciências) – Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2019.

Como discutido em Diaconis e Mosteller (1989), o problema dos aniversários é um ponto importante de partida no estudo de coincidências. Uma definição de coincidência é adotada, seguida de uma breve apresentação dos seus principais aspectos: entre eles a subjetividade e a surpresa. Na literatura são encontrados as principais variações do problema dos aniversários, que analisadas, fornecem as bases para um estudo do aspecto probabilístico das coincidências. Neste trabalho os eventos são associados a variáveis aleatórias independentes e igualmente distribuídas, ou seja, os problemas são estudados com o uso de técnicas de contagem. Desta forma, o problema clássico dos aniversários pode ser associado à coloração dos vértices de um grafo completo, assim como o problema do aniversariante pode ser associado à coloração dos vértices de um grafo estrela. A partir destas observações, é proposta a formulação do problema dos aniversários sujeito a restrições de uma rede de relacionamentos, que equivale a um problema da coloração própria dos vértices de um grafo simples, cuja topologia modela as restrições do problema original. O objetivo do trabalho é aplicar o conhecimento desenvolvido sobre polinômios cromáticos no estudo das coincidências associadas a problemas dos aniversários sujeito a restrições. Um resumo sobre teoria dos grafos e os principais resultados referentes a polinômios cromáticos são apresentados com o objetivos de dar clareza e consistência entre definições adotadas, teoremas utilizados e sua aplicação nos casos de árvores de Cayley, grafos Bollobás-Chung e redes sociais simples. A análise das coincidências se concentram na determinação do polinômio cromático e suas representações em diferentes bases, entre elas as formas: de potência, fatorial e de árvore. A decisão se um grafo pode ou não ser colorido de maneira própria com $k > 2$ cores é um problema de decisão NP-completo, portanto um objetivo secundário é analisar as limitações dos algoritmos existentes e dos sistemas disponíveis para cálculo de polinômios cromáticos, que possuem coeficientes inteiros e podem alcançar centenas de dígitos. Os cálculos e as simulações foram realizadas no Sage Mathematics Software (Linux Version 8.8). A conclusão demonstra que o conhecimento e as técnicas de polinômios cromáticos contribuem com a solução da generalização proposta do problema dos aniversários, sendo analisadas as limitações atuais quanto à eficácia dos algoritmos disponíveis e à capacidade computacional demandada pela topologia do grafo associado ao problema original.

Palavras-chaves: Coincidências. Problema dos aniversários. Polinômios cromáticos.

ABSTRACT

AMATO NETO, Luís. **Coincidences**: birthday problem and chromatic polynomials. 2020. 76 p. Dissertation (Master of Science) – School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, 2019.

As discussed in Diaconis e Mosteller (1989), the birthday problem is an important starting point in the study of coincidences. A definition of coincidence is adopted, followed by a brief presentation of its main aspects: subjectivity and surprise among them. In the literature, the main variations of the birthday problem are found, which analyzed, provide the basis for a study of the probabilistic aspect of coincidences. In this work the events are associated with independent and equally distributed random variables, that is, the problems are studied using counting techniques. Thus the classical birthday problem can be associated with the coloring of vertices of a complete graph, just as the birthmate problem can be associated with the coloring of the vertices of a star graph. From these observations, it is proposed a formulation of the birthday problem, subject to constraints of a relationships network, which is equivalent to a problem of proper coloring of the vertices of a simple graph, whose topology shapes the constraints of the original problem. The objective of this work is to apply the knowledge developed about chromatic polynomials in the study of coincidences associated with restricted birthday problems. A summary of graph theory and the main results concerning chromatic polynomials are presented with the purpose of providing clarity and consistency between adopted definitions, used theorems and their application in the case of Cayley trees, Bollobás-Chung graphs and simple social networks. The analysis of coincidences focuses on the determination of the chromatic polynomial and its representations in different bases, among them the power, factorial and tree bases. Whether or not a graph can be properly colored with $k > 2$ colors is an NP-complete decision problem, so a secondary objective is to analyze the limitations of existing algorithms and available systems for calculating chromatic polynomials, which have integer coefficients and can reach hundreds of digits. Calculations and simulations were performed in Sage Mathematics Software (Linux Version 8.8). The conclusion demonstrates that the knowledge and techniques of chromatic polynomials contribute to the solution of the proposed generalization of the birthdays problem, it is also analyzed the current limitations regarding the effectiveness of available algorithms and the computational capacity demanded by the graph topology associated with the original problem.

Keywords: Coincidences. Birthday problem. Chromatic polynomials.

LISTA DE FIGURAS

| | |
|--|----|
| Figura 2.1.1-Problema dos aniversários – Probabilidade de $k \geq 1$ coincidências | 18 |
| Figura 2.1.2-Problema do aniversariante - Probabilidade de $k \geq 1$ coincidências | 19 |
| Figura 2.2.1-Grafo completo e grafo estrela | 20 |
| Figura 2.3.1-Grafo circular: Polinômio cromático - Probabilidade de k coincidências | 21 |
| Figura 2.3.2-Grafo roda: Polinômio cromático - Probabilidade de k coincidências | 21 |
| Figura 2.3.3-Probabilidade de $k \geq 1$ coincidências x λ para diferentes grafos ou problemas | 21 |
| Figura 3.1.1-Exemplos de grafos | 23 |
| Figura 3.1.2-Exemplo de união de interseção de dois grafos G e H | 24 |
| Figura 3.1.3-Remoção de aresta e e remoção de vértice v | 25 |
| Figura 3.1.4-Identificação de vértices x e y e contração de aresta e | 25 |
| Figura 3.2.1-Exemplos de grafos G , G' e G'' | 28 |
| Figura 3.2.2-Exemplo de aplicação sucessiva do Teorema Fundamental e do Corolário | 29 |
| Figura 3.4.1-Exemplo de subgrafos de G gerados por clique separadora S | 33 |
| Figura 3.5.1-Exemplo de aplicação do Teorema dos Circuitos Quebrados | 37 |
| Figura 4.2.1-Probabilidade de y coincidência duplas | 40 |
| Figura 4.4.1-Probabilidade de quase aniversários com intervalo menor que k | 44 |
| Figura 5.0.1-Coefficientes na base de potência do polinômio cromático completo de $n = 30$ vértices | 47 |
| Figura 6.1.1-Árvore de Cayley de ordem 2: $q=3$, $q-1=2$ com 4 camadas | 49 |
| Figura 6.1.2-Variação dos coeficientes da base de potência | 51 |
| Figura 6.2.1-Grafos de Bollobas-Chung | 53 |
| Figura 6.2.2-Exemplos de grafo circular com 50 vértices $p=0,20$ ou 10 arestas aleatórias | 54 |
| Figura 6.2.3-Coefficientes dos polinômios cromáticos dos grafos circulares 50 vértices $p=0,20$ | 55 |
| Figura 6.2.4-Coefficientes dos polinômios cromáticos dos grafos circulares 50 vértices $p=0,20$ | 56 |
| Figura 6.2.5-Roteiro para geração de grafo circular com arestas aleatórias e polinômio cromático | 57 |
| Figura 6.3.1-Caso parceiras na mesa do jantar | 59 |

SUMÁRIO

| | | |
|----------|--|-----------|
| 1 | INTRODUÇÃO | 11 |
| 1.1 | Coincidências | 11 |
| 1.2 | Características das Coincidências | 12 |
| 1.3 | Problema de pesquisa | 15 |
| 2 | O PROBLEMA DOS ANIVERSÁRIOS | 17 |
| 2.1 | O problema clássico dos aniversários | 17 |
| 2.2 | O problema dos aniversários associado a grafos | 19 |
| 2.3 | Exemplos do problema dos aniversários em grafos | 20 |
| 3 | GRAFOS E POLINÔMIOS CROMÁTICOS | 23 |
| 3.1 | Elementos da teoria dos grafos | 23 |
| 3.2 | Coloração de vértices e polinômios cromáticos | 26 |
| 3.3 | Bases para representação de polinômios cromáticos | 30 |
| 3.4 | Teoremas importantes relativos ao cálculo de polinômios cromáticos | 32 |
| 3.5 | Interpretação dos coeficientes dos polinômios cromáticos | 34 |
| 4 | TIPOS DE PROBLEMA DOS ANIVERSÁRIOS | 38 |
| 4.1 | Problema clássico dos aniversários segundo von Mises | 38 |
| 4.2 | Problema clássico dos aniversários com múltiplos eventos | 40 |
| 4.3 | Outro problema clássico dos aniversários | 42 |
| 4.4 | Problema dos quase aniversários | 43 |
| 4.5 | Variações do problema dos aniversários | 45 |
| 5 | ALGORITMOS PARA CÁLCULO DE POLINÔMIOS CROMÁTICOS | 46 |
| 6 | CASOS DE APLICAÇÃO | 49 |
| 6.1 | Grafos Árvores de Cayley | 49 |
| 6.2 | Grafos de Bollobas-Chung | 53 |
| 6.3 | Grafos de Redes Sociais | 58 |
| 7 | CONCLUSÃO E RECOMENDAÇÕES | 60 |

| | |
|---|----|
| Referências | 62 |
| ANEXO A – ROTINA DE VON MISES (1939) | 65 |
| ANEXO B – ROTINA DE HOCKING E SCHWERTMAN(1986) | 66 |
| ANEXO C – ROTINA DE MCKINNEY (1966) | 68 |
| ANEXO D – ROTINA DE ÁRVORES DE CAYLEY | 70 |
| ANEXO E – ROTINA DE GRAFOS DE BOLLOBAS-CHUNG | 72 |
| ANEXO F – ROTINA DE REDES SOCIAIS | 74 |

1 INTRODUÇÃO

1.1 Coincidências

Coincidências fazem parte da nossa vida cotidiana. Os relatos de situações que incluem coincidências despertam nosso interesse, alimentam nosso imaginário e nos surpreendem. Podemos associar a elas mudanças no curso das nossas vidas: pessoas com quem passamos a conviver, lugares onde realizamos atividades de estudo ou de trabalho. Coincidências podem apontar para novas descobertas, assim como também para situações constrangedoras, como incidentes e acidentes.

Apesar de serem situações tão comuns, a formulação de problemas relativos às coincidências ainda é um desafio para a Ciência. Os diferentes campos do conhecimento têm abordado o assunto e contribuem para a compreensão de aspectos específicos do tema.

Segundo Diaconis e Mosteller (1989), coincidência pode ser definida como uma ocorrência surpreendente de eventos, percebidos como significativamente relacionados e sem uma causa aparente. Esta definição traz pelo menos cinco aspectos importantes: eventos próximos no tempo ou no espaço, percebidos como relacionados, surpreendentes, relevantes e sem causa aparente.

Diaconis e Mosteller abordam o tema coincidências de forma ampla e organizada a partir de métodos aplicáveis de estudo: estudos observacionais, experimentos, análise exploratória de dados e análise confirmatória de dados. Em cada método os autores fazem uma retrospectiva das principais contribuições e apontam caminhos a serem trilhados.

O primeiro método de estudo de coincidências, estudos observacionais, consiste em pesquisa de campo com a identificação de situações que cumpram requisitos definidos da pesquisa, coleta de dados e a busca de padrões ou de alguma forma de estruturação. Diaconis (1985) discute algumas vantagens e desvantagens deste método e cita como exemplos: observações não-controladas de casos médicos, testemunho de especialistas em processos legais. São citados outros métodos desenvolvidos mais recentemente, como grupos focais, historiografia e estudos de caso.

Os experimentos são conduzidos para testar um conjunto de fatores que podem ser controlados pelos investigadores, tais como estímulos e condicionantes, que se supõe significativos em relação aos resultados obtidos pelos experimentos. Os dados e resultados dos testes podem ser analisados, qualitativamente ou quantitativamente, por meio de técnicas estatísticas.

Quando dados estão disponíveis, sem prévio planejamento relativo a pesquisa, seu uso caracteriza o método de análise exploratória de dados e, tipicamente, tem como objetivo gerar hipóteses. Testes planejados destas hipóteses poderão confirmá-las, ou não, e são considerados como análise confirmatória de dados. Diaconis (1985) discute o uso, as limitações e os cuidados das técnicas estatísticas em ambas as análises de dados (exploratória e confirmatória).

Após sua publicação, Diaconis e Mosteller (1989) é citado praticamente em toda a literatura que aborda o tema coincidências, definida como ocorrência rara, surpreendente e sem causa aparente. A leitura de Peterson (1998), Hand (2014) e Mazur (2016), além da entrevista de Diaconis a Aldous (ALDOUS, 2013), permite afirmar que não houve um avanço significativo na formulação de problemas e abordagem proposta em Diaconis e Mosteller (1989), ou seja, o artigo é um adequado ponto de partida para este trabalho. A seguir estão apresentadas contribuições relevantes mais recentes sobre coincidências.

1.2 Características das Coincidências

Uma coincidência é uma associação de eventos feita por alguém. Estes eventos usualmente possuem uma sequência temporal e alguma proximidade espacial que permitem à pessoa relacioná-los. Se os eventos são significativos, relevantes e surpreendentes merecem atenção deste alguém. Se é possível atribuir causas conhecidas aos eventos, eles passam para a categoria de fato corriqueiro, caso contrário eles são denominados coincidências.

É possível agrupar estas ideias em dois grandes blocos: subjetividade e surpresa. O primeiro está relacionado à importância e relevância dos eventos para a pessoa, bem como sua capacidade de associar os eventos entre si e explicar seu nexos causal. O segundo trata da surpresa, que leva os eventos a serem percebidos como algo extraordinário.

O aspecto subjetivo é um assunto atual, tem como seminal o artigo de Tversky e Kahneman (1974), sendo que a trajetória destes dois importantes pesquisadores está condensada em Kahneman (2011). Em um brevíssimo resumo, as pessoas têm vieses cognitivos e sistematicamente falham na avaliação subjetiva da probabilidade dos eventos. Além disso, racionalidade não é suficiente para entender o valor atribuído às alternativas disponíveis e suas escolhas em um processo decisório. Em outras palavras, a teoria da decisão racional falha sistematicamente quando as decisões cotidianas são analisadas com mais detalhes.

As coincidências consideradas como surpresa incluem também um aspecto subjetivo, além do probabilístico. Por exemplo, em um estudo exploratório Griffiths e Tenenbaum (2007)

separa a análise das coincidências em dois grupos: as que são decorrentes basicamente de probabilidade, como uma mesma pessoa ganhar duas vezes na loteria, e aquelas que desafiam o conhecimento atual. Por exemplo, descobertas podem ter origem quando evidências desafiam a plausibilidade baseada no conhecimento prévio: intuições, avaliação de probabilidades e relações causais conhecidas. Dois dos cinco experimentos relatados utilizam o problema dos aniversários como estímulo. Estes experimentos indicam que a sensibilidade a coincidências pode despertar a curiosidade e é uma capacidade cognitiva que pode levar a descoberta de relações causais, tanto na Ciência como na vida cotidiana. A falta desta sensibilidade, se associada a crenças "irracionais", pode desencadear a propagação de crenças danosas às pessoas, por exemplo, a de que a vacinação é arriscada para crianças.

Outros exemplos de experimentos, Griffiths *et al.* (2018), complemento de Griffiths e Tenenbaum (2001), abordam a hipótese de que a aleatoriedade subjetiva é uma inferência estatística, ou seja, a apreensão da aleatoriedade pelas pessoas é uma inferência feita por elas sobre o processo que gerou os dados observados.

A raridade dos eventos associados à coincidência pode ser avaliada a partir da teoria das probabilidades, na tentativa de quantificar o quão ordinárias ou extraordinárias as coincidências podem ser, ou seja, esta avaliação pode contribuir para confrontar uma aproximação quantitativa com uma expectativa implícita na percepção de um evento como raro.

Dentre os diversos trabalhos recentes que tratam da aplicação da teoria das probabilidades ao estudo das coincidências ressaltamos dois que consideramos mais relevantes para este trabalho e que contribuem com respostas à seguinte pergunta: o que caracteriza uma coincidência?

O primeiro trabalho, Hand (2014), propõe um conjunto de "leis" que constituem "O Princípio da Improbabilidade":

- a) lei da inevitabilidade: algo deve acontecer;
- b) lei forte dos grandes números: com um número suficientemente grande de oportunidades, qualquer evento possível pode acontecer;
- c) lei da seleção: a probabilidade pode ser alterada para qualquer valor se o critério for escolhido após o evento;
- d) lei da alavancagem da probabilidade: uma pequena alteração das circunstâncias pode ter um enorme impacto nas probabilidades (não linearidade e sensibilidade às condições iniciais e de contorno);

e) lei do suficientemente próximo: eventos que são suficientemente similares podem ser percebidos como idênticos.

O livro traz inúmeros casos específicos, além de alguns genéricos, de aplicação desta abordagem para explicar a percepção de surpresa.

Em um breve resumo, a lei da inevitabilidade declara que a realidade está em inexorável movimento. Nosso conhecimento sobre este movimento é, e sempre será, limitado.

A lei forte dos grandes números reconhece que este movimento inexorável, em um período de tempo suficientemente longo, torna provável qualquer evento possível.

A lei da seleção nos lembra que uma "probabilidade" pode ser alterada para qualquer valor se a escolha do espaço amostral for feita após o evento considerado ter ocorrido.

A lei da alavancagem da probabilidade está relacionada à complexidade intrínseca ao movimento inexorável, ou seja, fenômenos complexos são tipicamente não-lineares e sensíveis às suas circunstâncias ou contexto (condições iniciais e de contorno). Esta lei também traz a questão de que estabelecer relações de causalidade, mesmo *a posteriori*, pode ser extremamente difícil, ou seja, a concorrência de eventos "sem causa aparente".

A lei do suficientemente próximo é semelhante a lei da seleção, quando a escolha do critério de eventos "idênticos" é tipicamente feita *a posteriori* por alguém que também considera na percepção a relevância dos eventos.

O segundo texto, Mazur (2016), também traz muitos exemplos e complementa esta caracterização das coincidências quando torna explícito dois outros aspectos complementares: a causalidade e o significado, como pode ser visto no trecho a seguir (p. 36):

Podemos buscar uma causa e procurar um significado. Causa e significado são duas coisas distintas. A causa de um evento é a razão principal pela qual o evento acontece. Há causas de um evento que não são determináveis, causas que são profundas demais para vislumbrarmos, e causas que são muito vagas para entendermos. Uma causa pode ter inúmeras camadas de entendimento.

O primeiro aspecto está relacionado ao repertório e à capacidade de alguém formular, ou não, uma cadeia de causalidade do evento e suas circunstâncias. Além disso, esta causalidade é estabelecida sempre *a posteriori*, pois, caso contrário, é uma previsão e contraria a definição adotada de coincidência.

O segundo aspecto considera a relevância e a carga emocional associadas ao significado atribuído por alguém à concorrência de eventos. Falk (1989) apresenta um exemplo de experimento, onde mostra que os participantes consideraram suas próprias coincidências mais

surpreendentes que as coincidências dos outros. Falk (2017) apresenta uma coleção de exemplos da associação de significado a coincidências, tanto nos textos acadêmicos quanto na literatura, notadamente de ficção. Para a autora há muito a ser melhor compreendido no aspecto psicológico das coincidências.

A primeira abordagem multidisciplinar das coincidências foi realizada por uma parceria entre Carl G. Jung e Wolfgang Pauli, iniciada em 1932 e interrompida pela morte de Pauli em 1958. Um dos resultados desta parceria foi a publicação de um livro com artigos relacionados, um de cada autor. Pauli (1955) associa as descobertas científicas a ideias arquetípicas relacionadas ao inconsciente coletivo, que é uma das ideias centrais da psicologia junguiana. Jung (1955) define sincronicidade como um princípio de conexões acausais, ou seja, traz a ideia de que para se entender a realidade são necessários dois princípios: o de causalidade e o de sincronicidade. O primeiro está associado aos fenômenos onde são identificadas relações de causa e efeito, enquanto o segundo está associado a fenômenos resultantes de eventos fortuitos e próximos no tempo e no espaço, ou seja, coincidências. Esta abordagem de associar fenômenos naturais e psique continua a ser desenvolvida, por exemplo, Cambray (2013) considera os conhecimentos mais recentes da Física e da Psicologia, e explora a associação entre sincronicidade e auto-organização de sistemas complexos.

Em resumo, a literatura é extensa em estudos observacionais e experimentos exploratórios sobre os diversos aspectos relacionados às coincidências, abrange vários campos do conhecimento que, aos poucos, vão sendo tratados em conjunto (multidisciplinaridade). Hand (2014) e Mazur (2016) aprofundam a caracterização das coincidências, entretanto muitas questões continuam em aberto, sendo a própria definição de coincidência uma delas.

Neste trabalho as coincidências são abordadas a partir do ponto de vista probabilístico, porém seus aspectos subjetivos não podem ser esquecidos na interpretação de dados da realidade. Por exemplo, o problema dos aniversários, na literatura também conhecido como paradoxo dos aniversários, é utilizado como exemplo didático em cursos de probabilidade e estatística porque causa "surpresa" nos alunos, sendo que os resultados usuais são relativos a um problema matematicamente bem definido e resolvido há décadas.

1.3 Problema de pesquisa

O objetivo geral deste trabalho é propor uma generalização do problema dos aniversários, quando a relação entre as pessoas envolvidas pode ser representada por uma rede de relacionamentos, e estudar as probabilidades de coincidências. Em outras palavras, em um modelo

ocupação de bolas (pessoas com datas de aniversário) distintas em caixas (datas do calendário) também distintas, a definição do espaço amostral obedece a restrições das caixas ocupadas que são consideradas como evento. Por exemplo, no problema clássico dos aniversários cada caixa individualmente pode ser ocupada e a variável aleatória é o número de caixas com s bolas; no problema clássico do aniversariante há somente dois tipos de ocupações de caixas: a do aniversariante e as outras, sendo a variável aleatória o número de bolas em cada tipo de caixa.

São tratados somente os casos mais simples destes problemas, que são aqueles onde as variáveis aleatórias são independentes e igualmente distribuídas, o que permite o estudo das probabilidades a partir das técnicas de contagem. A forma adotada para representação das restrições de relacionamento entre os participantes é a dos grafos simples e, naturalmente, coloração própria de vértices surge como técnica de contagem adequada ao problema dos aniversários. Outra vantagem desta escolha é o vasto conhecimento disponível desenvolvido no estudo dos problemas de coloração de grafos, tanto de polinômios quanto de números cromáticos.

Algumas famílias de grafos têm soluções conhecidas de seus polinômios cromáticos, tais como os grafos: completos, árvores, circulares, entre outros. O método de inclusão exclusão permite calcular o polinômio cromático de um grafo simples qualquer, porém decidir se um grafo pode ou não ser colorido de maneira própria com $k > 2$ cores é um problema NP-completo.

Para calcular a probabilidade das coincidências nos problemas dos aniversários é explorado o estudo dos coeficientes dos polinômios cromáticos representados em diferentes formas, sendo as mais adequadas a este estudo as formas: fatorial, de potência e de árvores. Assim como no cálculo de colorações próprias através de polinômios cromáticos, no caso mais geral do cálculo de probabilidade das coincidências é proposto um método de contração de arestas seguida de cálculo do polinômio cromático do grafo contraído.

Os casos estudados são: as árvores de Cayley, grafos de Bollobás-Chung e pequenas redes sociais. As árvores de Cayley são comuns em diversas áreas da Física, os grafos de Bollobás-Chung são os exemplos mais simples de uma rede com características "Small World" e as pequenas redes sociais apresentam em pequena escala as possibilidades e limitações da generalização proposta.

Polinômios cromáticos têm coeficientes inteiros e os valores de seus coeficientes podem ter centenas de dígitos. Estes números são arredondados em sistemas de cálculo tradicionais ou demandam uma programação específica para lidar com números inteiros com centenas de dígitos. A discussão das alternativas também é um objetivo secundário deste trabalho.

2 O PROBLEMA DOS ANIVERSÁRIOS

A primeira referência ao *problema dos aniversários* é atribuída a von Mises (1939). Mosteller (1962) formulou o problema como: qual é o menor número requerido de pessoas para que duas ou mais pessoas tenham a mesma data de aniversário com probabilidade $1/2$?

As simplificações usuais consistem em ignorar 29 de fevereiro como data de aniversário e que os outros 365 dias do ano são igualmente prováveis.

Muitas pessoas ficam surpresas com o resultado de que são necessárias "somente" 23 pessoas para que duas delas tenham a mesma data de aniversário com probabilidade de 50%.

Além do problema dos aniversários, Mosteller (1962) sugeriu um problema semelhante, denominado por ele de *problema do aniversariante*: dado um aniversariante, qual é o menor número requerido de pessoas para que uma outra pessoa tenha a mesma data de aniversário com probabilidade de $1/2$?

A partir das mesmas hipóteses, a resposta deste problema é 253 e também é surpreendente para um leigo no assunto.

A comparação da solução clássica destes problemas mostra que, no primeiro caso, a relação entre as pessoas envolvidas é de todas com todas, enquanto que, no segundo caso, é todas, menos uma, com uma única pessoa. Em outras palavras, a forma como as pessoas envolvidas são relacionadas definem diferentes problemas relacionados a aniversários. Portanto, é razoável propor uma generalização deste tipo de problema para diferentes configurações dos relacionamentos entre as pessoas, que é o ponto de partida para a formulação do *problema dos aniversários* utilizada neste trabalho.

2.1 O problema clássico dos aniversários

Feller (1968) apresenta o problema dos aniversários como um exemplo do problema da colocação aleatória de n bolas distintas (participantes e suas datas do aniversário) em λ caixas distintas (datas do ano). Um evento simples com n bolas ocorre quando pelo menos um par de bolas (participante com data do aniversário) é colocado na mesma caixa (data do ano), ou seja, pelo menos uma coincidência $k \geq 1$ (mesma data de aniversário). A probabilidade de cada caixa (data do ano) ser escolhida é a mesma, $1/\lambda = 1/365$.

No caso do *problema dos aniversários*, para calcular a probabilidade do evento de pelo menos uma coincidência é mais simples inicialmente calcular o seu complemento (nenhuma coincidência). A probabilidade de cada data do ano ser escolhida é a mesma, $1/\lambda = 1/365$.

Começando com duas pessoas: a probabilidade da segunda pessoa não ter a mesma data de aniversário da primeira é $364/365$. À cada pessoa adicionada a comparação é feita com todas as anteriores, logo a probabilidade da terceira pessoa não ter a mesma data de aniversário que a primeira e a segunda é $363/365$. Assim, sucessivamente, até a n -ésima data que tem probabilidade de $(365 - n + 1)/365$.

Uma data do ano pode receber qualquer número de datas de aniversário, logo são λ^n possibilidades no total. Portanto:

$$P(k = 0, n) = \underbrace{\frac{364}{365} \frac{363}{365} \dots \frac{365 - n}{365} \frac{365 - n + 1}{365}}_{n-1 \text{ termos}}, \quad n \geq 2. \quad (1)$$

Um produto deste tipo é denominado fatorial descendente, por conveniência, foi adotado como notação o símbolo de Pochhammer $\lambda^{(n)} = \lambda(\lambda - 1)(\lambda - 2) \dots (\lambda - n + 1)$, onde λ é número de caixas (datas do ano) e n é o número de bolas (participantes com datas de aniversários).

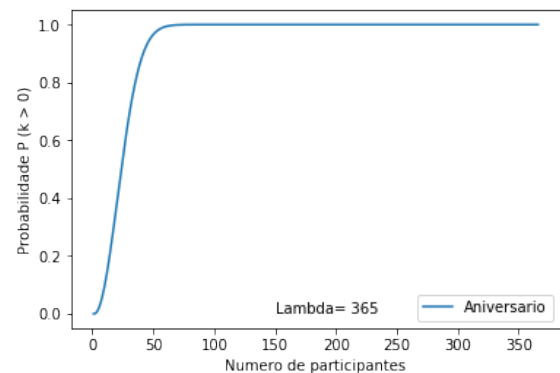
A figura 2.1.1 apresenta a probabilidade de pelo menos uma coincidência em função do número de participantes e $\lambda = 365$.

Figura 2.1.1 – Problema dos aniversários – Probabilidade de $k \geq 1$ coincidências

$$P(k = 0, n) = \frac{\lambda^{(n-1)}}{\lambda^{n-1}} = \frac{\lambda^{(n-1)} \cdot \lambda}{\lambda^{n-1} \cdot \lambda} = \frac{\lambda^{(n)}}{\lambda^n}$$

e

$$P(k \geq 1, n) = 1 - \frac{\lambda^{(n)}}{\lambda^n}$$



Fonte: Elaboração própria

Para facilitar as analogias entre diferentes configurações, por conveniência, neste trabalho o número de pessoas envolvidas nos problemas inclui todos os participantes, ou seja, são sempre n pessoas no total. No caso a seguir, o **aniversariante é incluído no total de participantes**.

Analogamente para o *problema do aniversariante*, começando com duas pessoas: a probabilidade da segunda pessoa não ter a mesma data de aniversário do aniversariante (primeira) é $364/365$. À cada pessoa adicionada a comparação é feita somente com o aniversariante, portanto a probabilidade da terceira pessoa não ter a mesma data de aniversário que a primeira

também é $364/365$. Assim, sucessivamente, até que a n -ésima data também tem probabilidade de $364/365$ de ser a mesma que a do aniversariante.

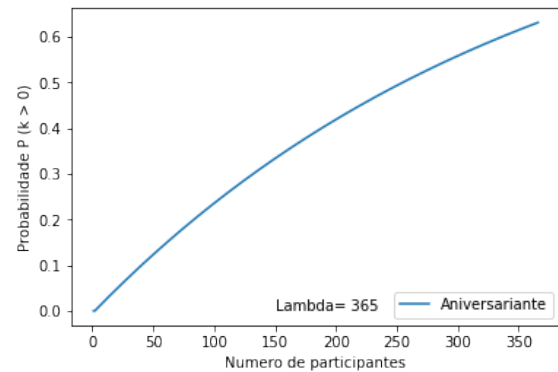
$$P(k = 0, n) = \underbrace{\frac{364}{365} \frac{364}{365} \cdots \frac{364}{365} \frac{364}{365}}_{n-1 \text{ termos}}, \quad n \geq 2. \quad (2)$$

Figura 2.1.2 – Problema do aniversariante - Probabilidade de $k \geq 1$ coincidências

$$P(k = 0, n) = \left(\frac{\lambda - 1}{\lambda}\right)^{n-1}$$

e

$$P(k \geq 1, n) = 1 - \left(\frac{\lambda - 1}{\lambda}\right)^{n-1}$$



Fonte: Elaboração própria

2.2 O problema dos aniversários associado a grafos

Falk (2014) formula o problema dos aniversários a partir da contagem do número de pares de comparações envolvidos no problema em relação ao número total de comparações possíveis e sugere que esta formulação pode dar origem a um modelo mais genérico do problema dos aniversários.

Apesar de publicado somente em 2016, desde 2009 Diaconis (2016) propôs que a solução do problema dos aniversários pode ser associado a um problema equivalente ao da coloração própria de grafos. Como definido em Read (1968), um polinômio cromático é o número de colorações próprias dos vértices de um grafo G e é denotado neste trabalho como $P_G(\lambda)$, onde λ é o número de cores.

Inicialmente, seja K_n o grafo completo de n vértices. Seu polinômio cromático é dado por

$$P_{K_n}(\lambda) = \lambda^{(n)}, \quad (3)$$

onde λ é número de cores. O número total de colorações possíveis de G é dado por λ^n , e se definirmos a probabilidade do grafo K_n ter coloração própria como:

$$P(k=0, n) = \frac{P_{K_n}(\lambda)}{\lambda^n} = \frac{\lambda^{(n)}}{\lambda^n}, \quad (4)$$

e daí

$$P(k \geq 1, n) = 1 - P(k=0, n) = 1 - \frac{\lambda^{(n)}}{\lambda^n}, \quad (5)$$

que é a solução do problema clássico dos aniversários.

Analogamente, o polinômio cromático do grafo estrela S_n de n vértices é dado por:

$$P_{S_n}(N) = \lambda(\lambda - 1)^{n-1}, \quad (6)$$

de forma que a probabilidade do grafo S_n ter coloração própria é

$$P(k=0, n) = \frac{P_{S_n}(\lambda)}{\lambda^n} = \frac{\lambda(\lambda - 1)^{n-1}}{\lambda^n} = \left(\frac{\lambda - 1}{\lambda}\right)^{n-1}, \quad (7)$$

então

$$P(k \geq 1, n) = 1 - P(k=0, n) = 1 - \left(\frac{\lambda - 1}{\lambda}\right)^{n-1}, \quad (8)$$

que é a solução do problema clássico do aniversariante.

Figura 2.2.1 – Grafo completo e grafo estrela



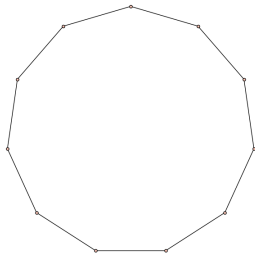
Fonte: Elaboração própria

2.3 Exemplos do problema dos aniversários em grafos

A literatura apresenta diversos casos de grafos cujos polinômios cromáticos têm solução conhecida Weisstein (2019). Os seguintes grafos foram selecionados como exemplo:

- Grafo circular ou n -gono: onde cada vértice é vizinho de apenas dois outros vértices, logo forma um ciclo, de onde resulta a denominação C_n , onde n é o número de vértices.
- Grafo roda: onde um vértice está conectado a todos os vértices de um grafo circular, neste trabalho é denominado W_n , onde n é o número de vértices.

Figura 2.3.1 – Grafo circular: Polinômio cromático - Probabilidade de k coincidências

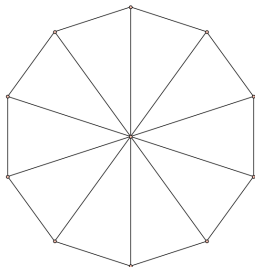


$$P(k=0, n) = \frac{(-1)^{n-1}(\lambda - 1) + (\lambda - 1)^{n-1}}{\lambda^n}$$

e

$$P(k \geq 1, n) = 1 - \frac{(-1)^{n-1}(\lambda - 1) + (\lambda - 1)^{n-1}}{\lambda^n}$$

Figura 2.3.2 – Grafo roda: Polinômio cromático - Probabilidade de k coincidências



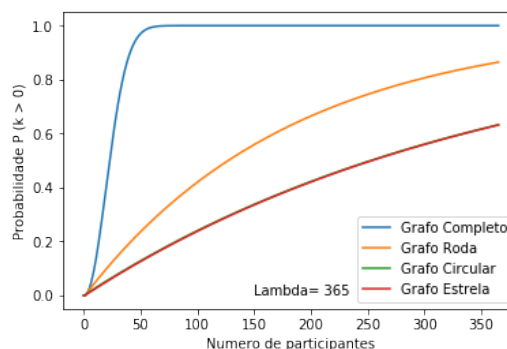
$$P(k=0, n) = \frac{\lambda[(\lambda - 2)^{n-1} - (-1)^n(\lambda - 2)]}{\lambda^n}$$

e

$$P(k \geq 1, n) = 1 - \frac{\lambda[(\lambda - 2)^{n-1} - (-1)^n(\lambda - 2)]}{\lambda^n}$$

A probabilidade de $k \geq 1$ coincidências referente aos diferentes problemas representados pelos grafos é apresentada na Figura 2.3.3 a seguir:

Figura 2.3.3 – Probabilidade de $k \geq 1$ coincidências x λ para diferentes grafos ou problemas



Fonte: Elaboração própria

A figura 2.3.3 mostra como as probabilidades variam fortemente com as regras de relacionamento dos participantes dos problemas dos aniversários em grafos. A partir desta constatação, este trabalho explora a generalização das regras de relacionamento entre participantes de problemas dos aniversários. Considerando a hipótese de eventos equiprováveis, modelar a probabilidade

de coincidências em redes de relacionamento se transforma em um problema de contagem em grafos, sendo que a coloração de vértices é a forma adotada neste trabalho.

A caracterização da rede é considerada um dado de entrada, ou seja, as informações para construção de um grafo representativo de alguma situação de interesse são disponíveis. A simulação das coincidências é sensível ao tamanho e complexidade da rede e, portanto, está limitada pela eficiência dos algoritmos e pela capacidade computacional disponível.

Além disso, são explorados os casos de coincidências múltiplas tratadas a partir do número de possibilidades geradas pela contração de arestas de uma coloração própria: uma a uma, no caso de coincidências simples; duas a duas, no caso de coincidências duplas; analogamente para coincidências de graus maiores. As quase coincidências podem ser tratadas pela soma das soluções referentes aos números de cores considerados coincidentes.

3 GRAFOS E POLINÔMIOS CROMÁTICOS

3.1 Elementos da teoria dos grafos

A apresentação de elementos da teoria dos grafos têm como objetivos: dar clareza, suportar a coerência e a consistência deste trabalho. Definições, nomenclatura, simbologia e teoremas são baseados em Feofiloff, Kohayakawa e Wakabayashi (2011) e Bondy e Murty (2008).

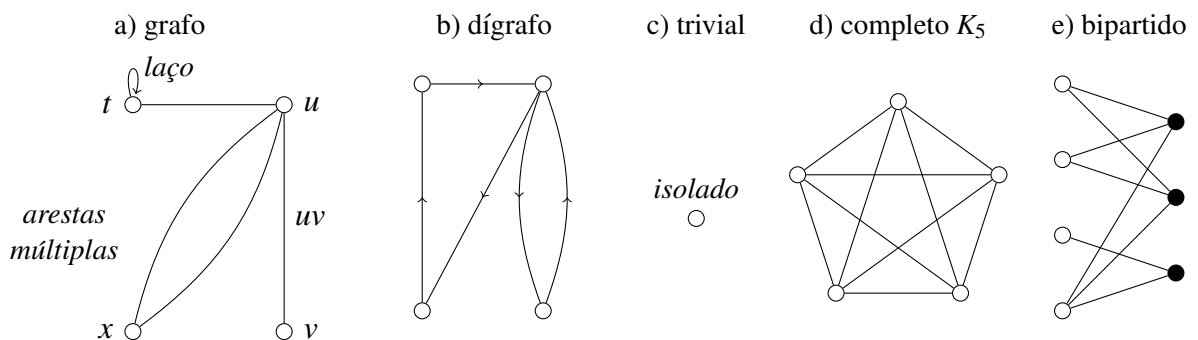
Um *grafo* G consiste em um conjunto de *vértices* $V(G)$ e um conjunto de *arestas* $A(G)$, onde cada aresta é definida por um par de vértices de $V(G)$, distintos ou não. Neste trabalho uma aresta é um par não-ordenado de vértices e G é denominado simplesmente *grafo*. Se cada aresta for definida por um par ordenado de vértices, G é denominado um *dígrafo*.

Se os vértices u e v são os extremos de uma aresta, representada por uv , então u e v são *adjacentes* ou *vizinhos*. Analogamente, se duas arestas têm um extremo (vértice) em comum, então as arestas são *adjacentes*. Arestas com os mesmos extremos (vértices) são chamadas *paralelas* ou *múltiplas*. Uma aresta com extremos iguais é chamada de *laço*.

Um grafo sem laços ou arestas múltiplas é denominado *simples*. Um grafo G é *nulo* se $V(G) = A(G) = \emptyset$. Um grafo G é *vazio* se $A(G) = \emptyset$. Um grafo com apenas um vértice e nenhuma aresta (*vértice isolado*) é chamado *trivial*.

A Figura 3.1.1 mostra alguns exemplos de grafos.

Figura 3.1.1 – Exemplos de grafos



Fonte: Elaboração própria

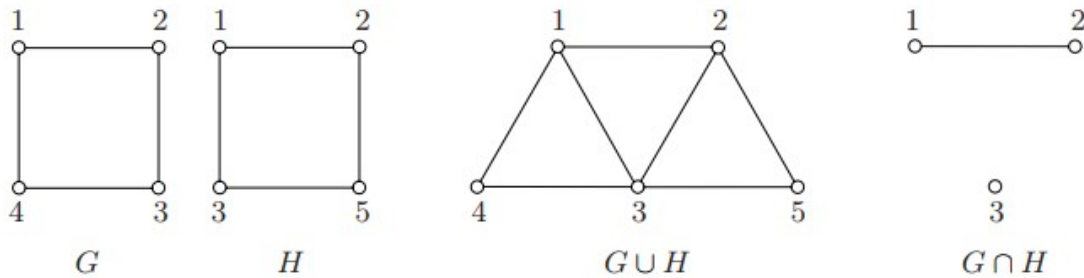
Dado um grafo G o grau de um vértice v é o número de arestas que incidem em v , denotado por $g_G(v)$. Um laço é contado como duas incidências. Um vértice que não tem arestas incidentes é denominado *isolado* e tem grau zero. Quando a referência ao grafo for evidente pelo contexto, o grau de um vértice v é denotado simplesmente por $g(v)$.

A *união* de dois grafos simples G e H é o grafo $G \cup H$ com $V(G \cup H) = V(G) \cup V(H)$ e $A(G \cup H) = A(G) \cup A(H)$.

A *interseção* de dois grafos simples G e H é o grafo $G \cap H$ com $V(G \cap H) = V(G) \cap V(H)$ e $A(G \cap H) = A(G) \cap A(H)$.

A Figura 3.1.2 mostra um exemplo de união e outro de interseção de dois grafos simples.

Figura 3.1.2 – Exemplo de união e interseção de dois grafos G e H



Fonte: Adaptado de Bondy e Murty, 2008

Um grafo simples G é *k-regular* se todos os seus vértices têm grau k ; G é denominado *regular* se é k -regular para algum k . Um grafo G é *bipartido* se $V(G)$ pode ser particionado em dois conjuntos X e Y tal que $X \cup Y = V(G)$ e $X \cap Y = \emptyset$ de modo que cada aresta de G tenha um extremo em X e outro em Y .

Um grafo simples G é *completo* se quaisquer dois de seus vértices são adjacentes. Um grafo completo com n vértices é denotado por K_n . O grafo K_3 é usualmente chamado de *triângulo*.

Dado um grafo simples G , o complemento de G , denotado por \overline{G} , é o grafo simples com os mesmos vértices $V(\overline{G}) = V(G)$, sendo que dois vértices são adjacentes em \overline{G} se e só se não são adjacentes em G , ou seja, $A(G) \cap A(\overline{G}) = \emptyset$ e $G \cup \overline{G}$ é completo.

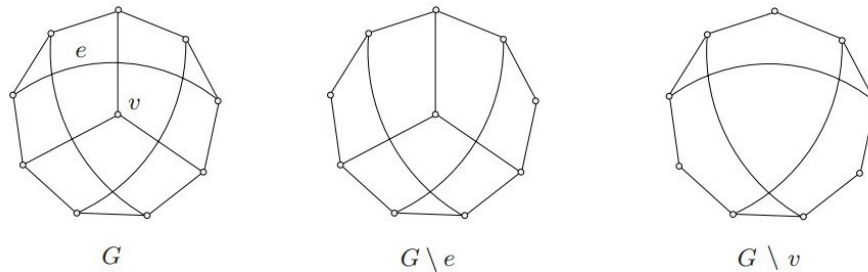
Um grafo H é um *subgrafo* do grafo G se $V(H) \subseteq V(G)$ e $A(H) \subseteq A(G)$ e é denotado por $H \subseteq G$. Analogamente, $G \supseteq H$, ou G é um *supergrafo* de H . Um subgrafo H de G é um *subgrafo gerador* (*spanning subgraph*) de G se $V(H) = V(G)$, ou seja, H é obtido pela remoção de arestas de G . Analogamente, S é um *supergrafo gerador* (*spanning supergraph*) de G se $V(S) = V(G)$, ou seja, S é obtido pela adição de arestas em G .

Dado um grafo G , seja X um conjunto de vértices com $\emptyset \subset X \subseteq V(G)$, então o subgrafo H de G *induzido* (ou *gerado*) por X é o subgrafo de G tal que $V(H) = X$ e $A(H)$ é o conjunto de todas as arestas de G que têm ambos os extremos em X , e é denotado por $G[X]$.

Um subgrafo pode ser obtido pela *remoção de uma aresta* e , sem alteração dos vértices, denotada por $G \setminus \{e\}$ ou simplesmente $G \setminus e$. Analogamente, pode ser obtido pela *remoção de*

um vértice v , e todas suas arestas incidentes, denotada por $G \setminus \{v\}$ ou simplesmente $G \setminus v$. A Figura 3.1.3 apresenta exemplos de ambos os casos.

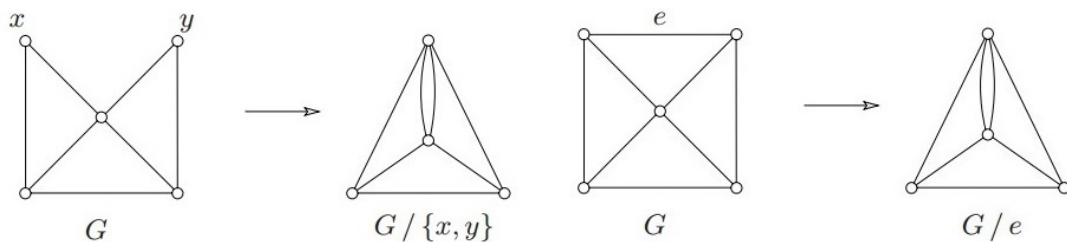
Figura 3.1.3 – Remoção de aresta e e remoção de vértice v



Fonte: Adaptado de Bondy e Murty, 2008

A *identificação* de dois vértices não adjacentes x e y de um grafo G é a sua substituição por um único vértice com todas as arestas incidentes, tanto em x como em y , denotada por $G/\{x,y\}$. Analogamente a *contração* de uma aresta e de um grafo G consiste na remoção da aresta e e na identificação dos extremos de e em um único vértice. Esta contração é denotada por G/e . A Figura 3.1.4 mostra um exemplo de cada um destes casos.

Figura 3.1.4 – Identificação de vértices x e y e contração de aresta e



Fonte: Adaptado de Bondy e Murty, 2008

Um *passeio* em um grafo G é uma sequência finita não vazia $P = (v_0; e_1; v_1; e_2; \dots; e_k; v_k)$, cujos termos são alternadamente vértices v_i e arestas e_i , tal que para todo i , $1 \leq i \leq k$, os extremos de e_i são v_{i-1} e v_i . Os vértices v_0 e v_k são a *origem* e o *término* do passeio P , os outros vértices são chamados *vértices internos* de P . O comprimento de P é o seu número de arestas, denotado por $\|P\|$. Um passeio é *fechado* se sua origem coincide com seu término e $\|P\| > 0$. Uma *trilha* é um passeio sem arestas repetidas. Um *caminho* é um passeio sem vértices repetidos. Uma trilha fechada cuja origem e vértices internos são todos distintos é um *circuito*, semelhante a um colar; um circuito de comprimento n é denotado por C_n .

Um grafo G é *conexo* se para todo par de vértices distintos u e v existe um caminho de u a v . Um grafo que não é conexo é dito *desconexo*. Os subgrafos conexos maximais de um grafo são chamados *componentes*.

Um grafo G que não contém circuitos é *acíclico*. Uma *árvore* é um grafo acíclico conexo. Uma *floresta* é um grafo acíclico não necessariamente conexo, ou seja, é um grafo cujos componentes são árvores.

Um subgrafo gerador de um grafo G é obtido pela remoção de arestas de G . Para um grafo conexo G , um subgrafo minimal conexo tem um único caminho entre quaisquer vértices u e v , pois do contrário existe um ciclo e existe uma aresta que pode ser removida mantendo a conectividade de G . Este argumento é a base para a prova do teorema 3.1.1, conforme Bondy e Murty (2008, Theorem 4.6 p. 106):

Teorema 3.1.1 *Um grafo é conexo se e somente contém uma árvore geradora.*

A *distância* entre dois vértices u e v de um grafo G é o comprimento de um caminho mais curto de u a v , denotada por $d_G(u, v)$ ou simplesmente $d(u, v)$. Quando não existir nenhum caminho de u a v , a distância é definida como infinita. A maior das distâncias entre quaisquer dois vértices do grafo G é o seu *diâmetro*, denotado por $diam(G)$. O comprimento de um menor circuito de um grafo G é sua *cintura*, denotada por $cint(G)$. O comprimento de um maior circuito de um grafo G é sua *circunferência*, denotada $circ(G)$. Se G é acíclico, sua cintura é definida como 0 e sua circunferência como infinita.

Uma *clique* em um grafo G é qualquer conjunto de vértices de G dois a dois adjacentes. Em outras palavras, $H \subseteq V(G)$ é uma clique do grafo G se o grafo induzido $G[X]$ é completo.

3.2 Coloração de vértices e polinômios cromáticos

Polinômios cromáticos foram introduzidos pelo matemático norte-americano George D. Birkhoff em 1912 na expectativa de solucionar a conjectura que, para a coloração de mapas (grafos planares), bastam 4 cores. Apesar deste caminho não ter se mostrado frutífero, uma vasta literatura foi desenvolvida sobre o assunto. Neste trabalho somente são necessários os fundamentos sobre polinômios cromáticos e as principais referências consideradas são: o artigo de Read (1968), os livros de Bondy e Murty (2008) e Dong, Koh e Teo (2005).

A *coloração* de um grafo G é o resultado de atribuir a cada vértice v uma das cores de um dado conjunto finito de cores C . Uma coloração é dita *própria* quando todos os vértices têm cores distintas das cores dos seus vizinhos. Uma coloração que não satisfaz esta restrição é dita

imprópria. Este trabalho, assim como em quase todos os casos de interesse, envolve somente colorações próprias, portanto é conveniente que o termo coloração seja referente à coloração própria. Se houver algum caso em que esta premissa não seja aplicável, então o critério utilizado estará explicitado.

Uma função de particular interesse, associada a um grafo G , é a que expressa as diferentes maneiras de colorir os vértices de G em função de um número especificado de cores, denotado historicamente por λ . Assim a função $P_G(\lambda)$ é o número de maneiras de colorir o grafo G com λ cores, sendo que nem todas as cores precisam ser utilizadas. Quando o contexto for evidente, $P_G(\lambda)$ é denotado simplesmente por $P(\lambda)$.

Na contagem das diferentes colorações, dois aspectos associados à rotação cíclica dos vértices e à diferenciação das cores devem ser previamente definidos. Para o objetivo deste trabalho os critérios adotados para contagem das diferentes colorações são:

- a) os vértices devem ser tratados com se fossem pontos fixos no espaço, ou seja, rotações cíclicas são contadas como colorações distintas;
- b) a cor atribuída a cada vértice é determinante, ou seja, cada cor atribuída a um vértice é contada como uma coloração distinta.

Os seguintes grafos simples ilustram a aplicação destes critérios e introduzem a função $P(\lambda)$:

- a) árvore de 3 vértices: se for adotada uma das λ cores para o vértice central, sobram $\lambda - 1$ cores para cada uma das pontas, logo $P(\lambda) = \lambda(\lambda - 1)^2$;
- b) triângulo (3 vértices): se for adotada uma das λ cores para um vértice qualquer, sobram $\lambda - 1$ cores para o segundo vértice e $(\lambda - 2)$ cores para o terceiro vértice, logo $P(\lambda) = \lambda(\lambda - 1)(\lambda - 2)$;
- c) grafo completo de n vértices: generalizando o caso do triângulo, se for adotada uma das λ cores para um vértice qualquer, sobram $\lambda - 1$ cores para o segundo vértice, $\lambda - 2$ cores para o terceiro vértice e assim sucessivamente, logo $P(\lambda) = \lambda(\lambda - 1)(\lambda - 2) \cdots (\lambda - n + 1)$;
- d) grafo vazio de n vértices: como cada vértice é isolado, pode ser adotada qualquer das λ cores para cada um deles, logo $P(\lambda) = \lambda^n$.

Pode-se observar que as funções $P(\lambda)$ destes casos são polinômios. A partir do teorema apresentado a seguir, é possível demonstrar que a função $P(\lambda)$ é polinomial para qualquer grafo G sem laços, daí a denominação de *polinômio cromático*.

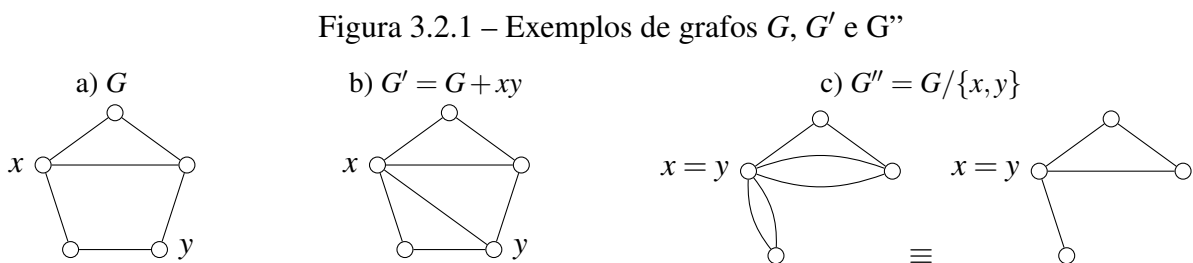
Seja um grafo G , não trivial e sem laços, com vértices x e y não adjacentes. Um exemplo é apresentado na Figura 3.2.1 a):

As colorações de G podem ser de dois tipos:

- a) aos vértices x e y são atribuídas *cores distintas*,
- b) aos vértices x e y são atribuídas as *mesmas cores*.

Se no primeiro caso as cores são distintas, então acrescentar uma aresta xy entre os vértices x e y do grafo G não altera a coloração destes casos.

Se no segundo caso as cores são as mesmas, então identificar os vértices x e y do grafo G não altera a coloração destes casos.



Fonte: Elaboração própria

Como estas alternativas são disjuntas, o número de colorações de um grafo G é a soma número das colorações do supergrafo G' (cores distintas para x e y) e do subgrafo G'' (mesmas cores para x e y). Esta é a motivação para o teorema que Read (1968) denomina de teorema fundamental dos polinômios cromáticos.

Teorema 3.2.1 (Fundamental) *Para um grafo G , sem laços, não trivial e com pelo menos um par de vértices não adjacentes x e y , $P_G(\lambda) = P_{G'}(\lambda) + P_{G''}(\lambda)$, onde $G' = G + xy$ e $G'' = G / \{x, y\}$.*

Os polinômios dos casos excluídos do teorema, a saber, o grafo trivial [$P_{G_n=1}(\lambda) = \lambda$] e grafos completos [$P_{K_n}(\lambda) = \lambda(\lambda - 1)(\lambda - 2) \cdots (\lambda - n + 1)$], completam a solução para qualquer grafo G sem laços.

O Teorema 3.2.1 pode ser aplicado sucessivamente a G' e G'' até que os grafos gerados não tenham mais nenhum par de vértices não adjacentes, ou seja, sejam completos. Utilizando um artifício gráfico proposto por Zykov (1949) *apud* Read (1968), um exemplo da sequência de aplicações sucessivas a um grafo é apresentada na Figura 3.2.2 a).

Figura 3.2.2 – Exemplo de aplicação sucessiva do Teorema Fundamental e do Corolário

The figure shows two methods of expanding a square graph (4 vertices, 4 edges) into a sum of complete graphs and paths. The left side (a) uses the Fundamental Theorem, and the right side (b) uses the Fundamental Corollary.

a) Teorema Fundamental

$$\begin{aligned}
 \square &= \square_{\text{diag}} + \text{L-shaped path} \\
 &= \square_{\text{diag}} + \underbrace{\triangle + \triangle}_{\text{two triangles}} + \text{vertical path} \\
 &= \square_{\text{diag}} + 2 \triangle + \text{vertical path}
 \end{aligned}$$

$\lambda^{(4)} + 2\lambda^{(3)} + \lambda^{(2)} =$

b) Corolário Fundamental

$$\begin{aligned}
 \square &= \square_{\text{diag}} - \triangle \\
 &= \text{two vertical paths} - \underbrace{\triangle + \triangle}_{\text{two triangles}} + \text{vertical path} \\
 &= \text{two vertical paths} - 2 \left\{ \text{diag 1} - \text{diag 2} \right\} + \text{vertical path} \\
 &= \text{two vertical paths} - 3 \text{ vertical path} + 3 \text{ vertical path} \\
 &= \begin{pmatrix} \circ & \circ & \circ \\ \circ & - & \circ \\ \circ & \circ & \circ \end{pmatrix} - 3 \begin{pmatrix} \circ & \circ \\ \circ & - \\ \circ & \circ \end{pmatrix} + 3 \begin{pmatrix} \circ & \circ \\ \circ & - \end{pmatrix} \\
 &= \begin{matrix} \circ & \circ & & \circ & \circ & \circ & & \circ \\ & \circ & \circ & & \circ & & \circ & & \circ \\ & & \circ & \circ & & \circ & & \circ & & \circ \\ & & & \circ & \circ & & \circ & & \circ & & \circ \\ & & & & \circ & & & \circ & & \circ & & \circ \end{matrix} - 4 \begin{matrix} \circ & \circ & \circ \\ \circ & - & \circ \\ \circ & \circ & \circ \end{matrix} + 6 \begin{matrix} \circ & \circ \\ \circ & - \end{matrix} - 3 \circ
 \end{aligned}$$

$\lambda^4 - 4\lambda^3 + 6\lambda^2 - 3\lambda$

Fonte: Adaptado de Read (1968)

A aplicação sucessiva do Teorema Fundamental resulta que o polinômio cromático de G é uma soma de polinômios cromáticos de grafos completos. Para expressar polinômios cromáticos de grafos completos é conveniente utilizar o símbolo de Pochhammer

$$\lambda^{(n)} = \lambda(\lambda - 1)(\lambda - 2) \dots (\lambda - n + 1), \tag{9}$$

onde n é número de vértices de G . Portanto, para grafos completos $P_{K_n}(\lambda) = \lambda^{(n)}$.

O polinômio cromático expresso como a soma dos polinômios resultantes da aplicação sucessiva do Teorema 3.2.1 resulta em

$$P_G(\lambda) = \sum_{i=1}^n a_i \lambda^{(i)}, \tag{10}$$

onde n é número de vértices de G e a_i são números inteiros com $a_n = 1$, pois o grafo G dá origem a um único grafo completo K_n , todos os outros têm menor número de vértices até chegar ao grafo trivial. Esta forma de $P_G(\lambda)$ é denominada *forma fatorial*.

O teorema 3.2.1 pode ser reescrito na forma de remoção de arestas da seguinte forma:

Corolário 3.2.2 (Fundamental) *Para um grafo G , sem laços, não trivial e com pelo menos uma aresta entre os vértices x e y , $P_{G'}(\lambda) = P_G(\lambda) - P_{G''}(\lambda)$, onde $G' = G + xy$ e $G'' = G/\{x, y\}$.*

O corolário 3.2.2 pode ser aplicado sucessivamente a G e G'' até que os grafos gerados não tenham mais nenhuma aresta, ou seja, sejam vazios. A Figura 3.2.2 b) apresenta o mesmo exemplo com a sequência de aplicações sucessivas do corolário a um grafo.

A aplicação sucessiva do corolário 3.2.2 resulta que o polinômio cromático de G é uma soma de polinômios cromáticos de grafos vazios,

$$P_G(\lambda) = \sum_{i=1}^n b_i \lambda^i, \quad (11)$$

onde n é número de vértices de G e b_i são números inteiros com $b_n = 1$, pois o grafo G dá origem a um único grafo vazio de n vértices, todos os outros têm menor número de vértices até chegar ao grafo trivial. Esta forma de $P_G(\lambda)$ é denominada *forma de potência*.

Assim fica demonstrado que a função $P_G(\lambda)$ é um polinômio para qualquer grafo G simples.

Vale ressaltar que: o problema de associar grafos a um dado polinômio é um problema em aberto, ou seja, é conhecido como associar um grafo a um polinômio cromático, porém o inverso é um problema em aberto.

3.3 Bases para representação de polinômios cromáticos

Independentemente do método de cálculo, o polinômio cromático de um grafo é único, portanto uma forma pode ser transformada em qualquer outra equivalente. A motivação para utilizar estas transformações é que os coeficientes associados a cada forma podem ser interpretados e revelar características interessantes associadas à coloração dos vértices dos grafos, sendo este um dos objetivos deste trabalho.

Um polinômio cromático de grafo conexo que é obtido por aplicações sucessivas do teorema 3.2.1 resulta em uma soma ponderada por números inteiros de polinômios de grafos completos, desde o grau n (número de vértices) até grau que representa o menor número cores que permite a coloração própria do grafo. Este grau é o número cromático do grafo, pois um número de cores inferior a este grau faz com que os valores de todos os polinômios dos grafos completos sejam nulos.

Analogamente, a aplicação sucessiva do corolário 3.2.2 resulta em uma soma ponderada por números inteiros de polinômios de grafos vazios, desde o grau n até 1. Um polinômio

cromático tem termo independente igual a zero, pois do contrário, para número de cores $\lambda = 0$ o valor de $P(\lambda)$ seria diferente de zero, ou seja, haveria colorações próprias com zero cores, o que é um absurdo. Neste caso, o número cromático é o menor inteiro positivo onde o polinômio cromático é positivo.

Uma alternativa de cálculo a partir da aplicação sucessiva do corolário 3.2.2 é a remoção de arestas de forma que os subgrafos gerados sejam árvores, o que reduz o número de remoções de arestas em relação à alternativa de grafos vazios. Como todas as árvores com k vértices têm o mesmo polinômio cromático, então um polinômio cromático pode ser representado pela soma ponderada de n polinômios na forma de $P(\lambda) = \lambda(\lambda - 1)^{k-1}; k = 1, 2, \dots, n$, ou seja, este conjunto é uma base para polinômios cromáticos (LOERINC, 1980).

Chao e Whitehead (1978) e Loerinc (1980) mostram as equações que relacionam as formas fatorial, de potência e de árvore:

a) forma fatorial para forma de potência

$$\lambda^{(n)} = \sum_{k=1}^n (-1)^{k+n} s(n, k) \lambda^k \quad (12)$$

onde: $s(n, k)$ é número de Stirling de 1ª espécie,

com: $s(n, 0) = 0$ para $n > 0$ e $s(n+1, k) = s(n, k-1) - n s(n, k)$ para $0 < k \leq n$.

Por exemplo, para $n=5$: $P(K_5) = \lambda^{(5)} = \lambda^5 - 10\lambda^4 + 35\lambda^3 - 50\lambda^2 + 24\lambda$

$\Leftrightarrow s(5, 5) = 1, s(5, 4) = 10, s(5, 3) = 35, s(5, 2) = 50, s(5, 1) = 24$.

b) forma de potência para forma fatorial

$$\lambda^n = \sum_{k=1}^n S(n, k) \lambda^{(k)} \quad (13)$$

onde: $S(n, k)$ é número de Stirling de 2ª espécie,

com: $S(n, 0) = 0$ para $n > 0$ e $S(n+1, k) = S(n, k-1) + k S(n, k)$ para $0 < k \leq n$.

Por exemplo, para $n=5$: $P(Vazios) = \lambda^5 = \lambda^{(5)} + 10\lambda^{(4)} + 25\lambda^{(3)} + 15\lambda^{(2)} + 1\lambda^{(1)}$

$\Leftrightarrow s(5, 5) = 1, s(5, 4) = 10, s(5, 3) = 25, s(5, 2) = 15, s(5, 1) = 1$.

c) forma de árvore para forma de potência

$$\lambda(\lambda - 1)^{n-1} = \sum_{k=1}^n (-1)^{n+k} t(n, k) \lambda^k, \text{ para } n \geq 1 \quad (14)$$

onde: $t(n, k) = \binom{n-1}{k-1}$.

Por exemplo, para $n=5$: $P(T_5) = \lambda(\lambda - 1)^4 = \lambda^5 - 4\lambda^4 + 6\lambda^3 + 4\lambda^2 + 1\lambda^1$

$\Leftrightarrow \binom{4}{4} = 1, \binom{4}{3} = 4, \binom{4}{2} = 6, \binom{4}{1} = 4, \binom{4}{0} = 1$.

d) forma de potência para forma de árvore

$$\lambda^n = \sum_{k=1}^n T(n,k) \lambda (\lambda - 1)^{n-1}, \text{ para } n \geq 1 \quad (15)$$

onde: $T(n,k) = \binom{n-1}{k-1}$.

Por exemplo, para $n=5$:

$$P(\text{Vazios}) = \lambda^5 = 1 \lambda (\lambda - 1)^4 + 4 \lambda (\lambda - 1)^3 + 6 \lambda (\lambda - 1)^2 + 4 \lambda (\lambda - 1)^1 + 1 \lambda (\lambda - 1)^0$$

$$\Leftrightarrow \binom{4}{4} = 1, \binom{4}{3} = 4, \binom{4}{2} = 6, \binom{4}{1} = 4, \binom{4}{0} = 1.$$

No caso dos itens c) e d) Loerinc (1980) apresenta somatória desde $k=0$, associado ao termo independente que é sempre nulo, e utiliza como definição $\binom{n-1}{-1} = 0$.

3.4 Teoremas importantes relativos ao cálculo de polinômios cromáticos

O Teorema Fundamental permite calcular o polinômio de um grafo, porém é fácil perceber que, com o aumento do número de vértices e de arestas, o número de polinômios gerados pela remoção / contração ou pela adição / contração cresce exponencialmente, o que está associado a característica deste problema ser NP-Completo.

As diferentes abordagens de algoritmos procuram identificar sequências de vértices e arestas que minimizem o número de polinômios gerados e que facilitem seu cálculo. Os seguintes teoremas relativos aos polinômios cromáticos contribuem de forma importante nas heurísticas dos algoritmos disponíveis na literatura.

Como pela definição os componentes são conexos e disjuntos entre si, a coloração de cada um é independente dos outros e, portanto, o número de colorações de G é o produto do número de colorações de seus componentes.

Teorema 3.4.1 *Se um grafo G tem k componentes conexas G_1, G_2, \dots, G_k então*

$$P_G(\lambda) = P_{G_1}(\lambda) \cdot P_{G_2}(\lambda) \cdots P_{G_k}(\lambda).$$

Em uma árvore existe somente um caminho entre dois vértices quaisquer, logo λ cores podem ser atribuídas um vértice qualquer e $\lambda - 1$ a cada um de seus vizinhos, assim sucessivamente a cada vizinho $\lambda - 1$ podem ser atribuídas, logo:

Teorema 3.4.2 *Se um grafo T_n é uma árvore com n vértices então*

$$P_{T_n}(\lambda) = \lambda (\lambda - 1)^{n-1}.$$

Para um grafo circular C_n com n vértices, a aplicação do corolário 3.2.2 resulta:

$$P_{C_n}(\lambda) = P_{T_n}(\lambda) - P_{C_{n-1}}(\lambda), \text{ onde } P_{T_n}(\lambda) = \lambda(\lambda - 1)^{n-1} \text{ e } P_{C_2} = \lambda(\lambda - 1).$$

A aplicação sucessiva desde C_n até C_2 resulta que $P_{C_n}(\lambda)$ é a soma de uma progressão geométrica de $n - 1$ termos com $a_0 = \lambda(\lambda - 1)^{n-1}$ e razão $r = -1/(\lambda - 1)$, ou seja:

$$P_{C_n}(\lambda) = \frac{a_0(1 - r^{n-1})}{1 - r} = \frac{\lambda(\lambda - 1)^{n-1}(1 - (-1/(\lambda - 1))^{n-1})}{1 - (-1/(\lambda - 1))} = (\lambda - 1)^n(1 - (-1/(\lambda - 1))^n),$$

então:

Teorema 3.4.3 *Se um grafo C_n é circular com n vértices então*

$$P_{C_n}(\lambda) = (\lambda - 1)^n + (-1)^n(\lambda - 1).$$

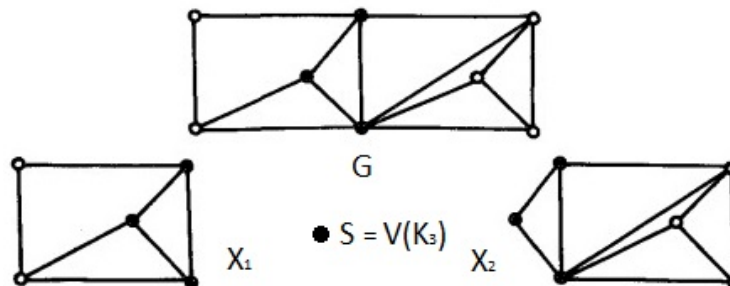
Para o grafo conexo G , um conjunto de vértices $S \subset V(G)$ é separador (ou conjunto de corte) se o grafo $G \setminus S$, resultante da remoção dos vértices de S de G , é desconexo. Além disso, se S é uma clique em G , então S é uma *clique separadora* em G . Neste caso o grafo G pode ser separado em subgrafos cuja superposição é $G[S]$, ou seja, possuem $G[S]$ em comum (Read (1968, Theorem 3 p. 59) e Dong, Koh e Teo (2005, Theorem 1.3.2 p. 8)). Como $G[S]$ é completo, então se for atribuída uma coloração a $G[S]$ a coloração dos vértices restantes de cada subgrafo são independentes. Logo:

Teorema 3.4.4 *Se a "superposição" de dois grafos X e Y é um grafo completo com k vértices, então o polinômio cromático formado pela junção de X e Y é dado por:*

$$P_G(\lambda) = \frac{P_X(\lambda) \cdot P_Y(\lambda)}{P_{K_k}(\lambda)} = \frac{P_X(\lambda) \cdot P_Y(\lambda)}{\lambda^{(k)}}$$

A Figura 3.4.1 apresenta um exemplo do Teorema 3.4.4 com $k=2$ e $S = V(K_3)$.

Figura 3.4.1 – Exemplo de subgrafos de G gerados por clique separadora S



Fonte: Adaptado de Read (1968)

3.5 Interpretação dos coeficientes dos polinômios cromáticos

Read (1968) mostra que é possível interpretar os coeficientes dos polinômios cromáticos, sendo estas interpretações valiosas para o estudo de coincidências.

Para um grafo G com n vértices, seu polinômio cromático $P_G(\lambda)$ tem as seguintes propriedades:

- O grau de $P(\lambda)$ é n .
- O coeficiente de λ^n em $P(\lambda)$ é 1.
- $P(\lambda)$ não tem termo constante.
- Os termos de $P(\lambda)$ na base de potência alternam de sinal.
- O valor absoluto do coeficiente de λ^{n-1} é o número de arestas.
- Para G conexo, $P(\lambda) \leq \lambda(\lambda - 1)^{n-1}$, onde λ é inteiro positivo.
- G é uma árvore se e somente se $P(\lambda) = \lambda(\lambda - 1)^{n-1}$.
- Se G é conexo, então o valor absoluto do coeficiente de λ^r em $P(\lambda)$ não é menor que $\binom{n-1}{r-1}$.
- O menor valor de r , cujo coeficiente de λ^r é diferente de zero, é o número de componentes de G .

A forma fatorial permite a seguinte interpretação. Para $\lambda = r$, $P_G(r)$ é o número de colorações próprias com exatamente r cores, quaisquer que sejam as r cores de um conjunto de λ cores, ou seja, é o número de partições do conjunto de vértices de G em r subconjuntos disjuntos de vértices, tal que não há vértices vizinhos em cada subconjunto. Considerando as permutações das r cores, resulta que são $r!P_G(r)$ colorações próprias com exatas r cores.

O polinômio cromático $P_G(\lambda)$ é o número de colorações próprias com λ cores, mas não é necessário o uso de todas elas. Este fato permite reconstruir o polinômio cromático através da soma de todas as colorações próprias dos r subconjuntos disjuntos de r cores, $r = 1, 2, \dots, \lambda$. Como o número de combinações das λ cores de r maneiras é $\binom{\lambda}{r}$, então:

$$P_G(\lambda) = \sum_{r=1}^{\lambda} \binom{\lambda}{r} r! P_G(r) = \sum_{r=1}^{\lambda} P_G(r) \lambda^{(r)}$$

Esta é a justificativa para o seguinte teorema:

Teorema 3.5.1 *O coeficiente de $\lambda^{(r)}$ da forma fatorial de $P_G(\lambda)$ é o número de maneiras de colorir o grafo G com exatas r cores com colorações indiferentes.*

Uma aplicação importante deste teorema é que o menor expoente do polinômio cromático na forma fatorial é o menor número de cores que permite colorações próprias, logo é o número cromático $\chi(G)$.

A interpretação dos coeficientes do polinômio cromático na forma de potência foi proposta por Whitney (1932), a partir do princípio da inclusão exclusão aplicada à propriedade coloração.

Dado um grafo simples conexo G com n vértices e m arestas, o número de colorações próprias pode ser calculado partindo do número total de colorações (λ^n) subtraindo o número de colorações impróprias.

Para uma coloração qualquer de G , se forem removidas as arestas cujos vértices têm cores distintas, o subgrafo resultante tem k componentes, cada um com uma única cor, logo são λ^k colorações ditas altamente impróprias. De outra forma, para cada coloração própria ou imprópria há uma coloração altamente imprópria. Para coloração própria este subgrafo de coloração altamente imprópria é um grafo vazio.

Seja $N(k, r)$ o número de subgrafos com k componentes e r arestas. Utilizando o método da inclusão exclusão, o número de colorações próprias pode ser calculado por:

$$P_G(\lambda) = \lambda^n - \sum_p N(p, 1)\lambda^p + \sum_p N(p, 2)\lambda^p + \dots (-1)^m \sum_p N(p, m)\lambda^p$$

Como o subgrafo com nenhuma aresta é o grafo vazio, então $N(n, 0) = 1$. Agrupando os termos:

$$P_G(\lambda) = \sum_{r=0}^m \sum_{p=1}^n (-1)^r N(p, r)\lambda^p = \sum_{p=1}^n \left\{ \sum_{r=0}^m (-1)^r N(p, r) \right\} \lambda^p \quad (16)$$

De onde decorre:

Teorema 3.5.2 *O coeficiente de λ^p da forma de potência de $P_G(\lambda)$ é $\sum_{r=0}^m (-1)^r N(p, r)$, onde $N(p, r)$ é o número de subgrafos de G com p componentes e r arestas, e m o número de arestas de G .*

Whitney (1932) mostrou que, para cada conjunto de subgrafos de p componentes, várias parcelas se anulam, o que reduz o esforço de contagem dos $N(p, r)$ subgrafos. O argumento se baseia na constatação de que, fixado o número de componentes p , para um componente do subgrafo $H_{p, r+1}$ com $r+1$ arestas que contém pelo menos um circuito, existe um subgrafo $H_{p, r}$ com r arestas que se obtém pela remoção de uma aresta do circuito, sendo que o primeiro contribui com o sinal $(-1)^{r+1}$ e o segundo com $(-1)^r$, ou seja, se anulam.

Para assegurar a correta contagem do número de subgrafos $N(p, r)$, Whitney (1932) propõe o seguinte algoritmo:

- a) adotar uma ordem qualquer para os vértices do grafo G ;
- b) listar os circuitos identificando os vértices na ordem definida em a);
- c) para cada circuito, construir um "circuito quebrado" removendo a última aresta deste circuito;
- d) adotar uma ordem qualquer para os circuitos quebrados;
- e) seja Q_i cada um dos q circuitos quebrados, $i = 1, 2, \dots, q$;
- f) dividir todos os subgrafos de G em $q + 1$ conjuntos, alguns dos quais podem ser vazios, sendo que o conjunto S_1 contém todos os subgrafos que contém todas as arestas de Q_1 , S_2 contém todos os subgrafos que contém todas as arestas de Q_2 e não estão contidos em S_1 , S_3 contém todos os subgrafos que contém todas as arestas de Q_3 e não estão contidos em $S_1 \cup S_2$, assim sucessivamente até S_{q+1} que contém todos os subgrafos que não contém nenhum dos circuitos quebrados.

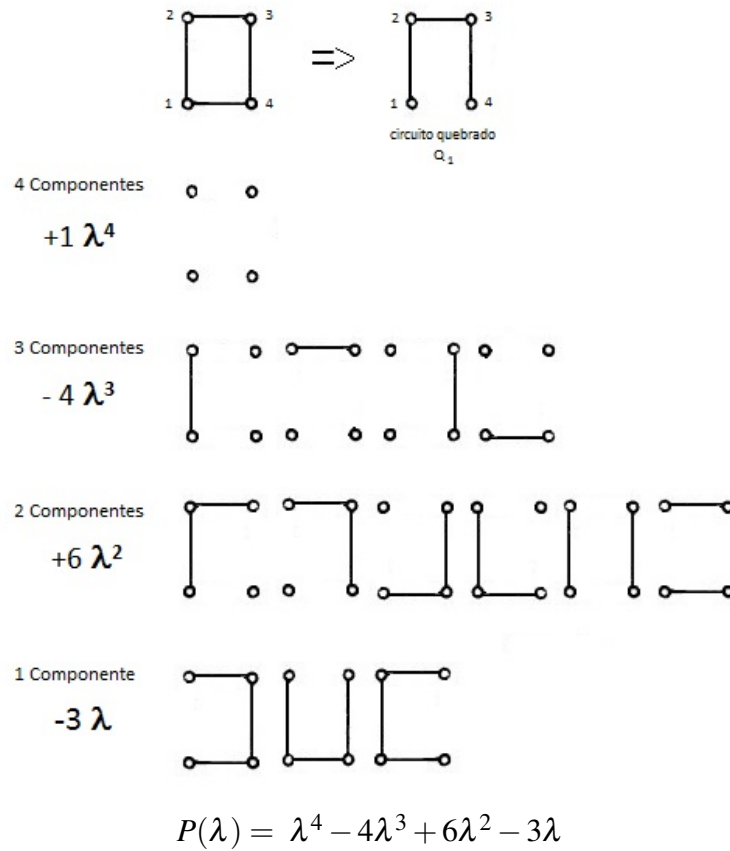
Para demonstrar que as parcelas na contagem dos $N(p, r)$ se anulam em cada um dos conjuntos S_i , $i = 1, 2, \dots, q$ é utilizado o seguinte argumento. Considere a aresta a_1 que foi removida do circuito para construir o circuito quebrado Q_1 . A cada subgrafo de S_1 que contém a_1 corresponde um subgrafo que não contém a_1 , de outra forma, por construção existe um subgrafo em S_1 que corresponde ao subgrafo que não contém a_1 ao qual se adiciona a_1 . Como a diferença do número de arestas entre este par é de uma aresta, a contribuição na contagem dos $N(p, r)$ subgrafos do termo de λ^p é nula. Argumento semelhante vale para todos os outros conjuntos porque sucessivamente eles não incluem os circuitos quebrados anteriores.

Portanto o único conjunto de subgrafos que contribui na contagem de $N(p, r)$ é S_{q+1} , ou seja, os subgrafos que não contém nenhum circuito quebrado do grafo G .

Teorema 3.5.3 *O coeficiente de λ^p da forma de potência de $P_G(\lambda)$ é $\sum_{r=0}^m (-1)^r N(p, r)$, onde $N(p, r)$ é o número de subgrafos de G com p componentes e r arestas, que não contém nenhum circuito quebrado de G , sendo m o número de arestas de G .*

A figura 3.5.1 apresenta a aplicação deste teorema ao mesmo exemplo utilizado na figura 3.2.2. Este exemplo é bastante simples porque há somente um circuito quebrado.

Figura 3.5.1 – Exemplo de aplicação do Teorema dos Circuitos Quebrados



Fonte: Elaboração própria

O ponto importante é que a análise da figura mostra que os subgrafos gerados, associados a cada coeficiente de λ^p , indicam um conjunto de arestas com coloração altamente imprópria (uma única cor) que podem ser contraídas para o cálculo das coincidências no polinômio cromático do grafo G .

4 TIPOS DE PROBLEMA DOS ANIVERSÁRIOS

No seu estudo sobre coincidências, Diaconis e Mosteller (1989) abordam a questão de modelos genéricos e modelos especializados. O primeiro tipo são aqueles que têm uma formulação matemática que pode ser aplicada a diferentes situações, por exemplo: o modelo de ocupação de caixas por bolas, enquanto que o segundo tipo trata de caso específico. A pertinência de um modelo a um grupo não é definitiva, pois a aplicação de modelo genérico pode identificar a necessidade de modelo específico, bem como um modelo específico pode ter sua aplicação ampliada e dar origem a um modelo genérico.

Os modelos genéricos do problema dos aniversários são classificados em quatro grupos:

- a) problema clássico: semelhante a Mosteller (1962),
- b) múltiplos eventos: coincidências de três ou mais elementos,
- c) quase aniversários: intervalo ampliado para que o evento seja considerado uma coincidência,
- e) múltiplas categorias: coincidências quando cada elemento tem mais que uma categoria.

4.1 Problema clássico dos aniversários segundo von Mises

Feller (1968) atribui a primeira referência ao *problema dos aniversários* a von Mises (1939). O modelo proposto por von Mises é de uma ocupação de λ caixas distintas por n bolas distintas, em relação ao definido na seção 2.1, são λ cores atribuídas a n vértices distintos.

Sejam as variáveis aleatórias independentes e igualmente distribuídas $X_s : s = 0, 1, 2, \dots, k$ a quantidade de caixas com s bolas, sendo X_0 a quantidade de caixas vazias.

A distribuição de probabilidades da ocupação de λ caixas por n bolas é definida por $P(x_0, x_1, x_2, \dots, x_k)$, quando se conhece os valores x_i que indicam quantas caixas estão vazias, com uma, com duas, ..., com k bolas, sujeitos às seguintes restrições:

- a) número de caixas: $x_0 + x_1 + x_2 + \dots + x_k = \lambda$
- b) número de bolas: $1x_1 + 2x_2 + \dots + kx_k = n$

Estas restrições significam que, por exemplo, para um valor de s maior que $k/2$, x_s só pode assumir os valores 0 ou 1.

A probabilidade de uma ocupação $x_0, x_1, x_2, \dots, x_k$ pode ser encontrada determinando o número de diferentes arranjos que são possíveis com uma dada ocupação em relação ao número

total de ocupações possíveis. A tarefa relativa ao número de diferentes arranjos possíveis pode ser dividida em três fatores:

a) Arranjo das λ caixas em $k + 1$ grupos de caixas (vazias, 1 bola, 2 bolas, ..., k bolas):

$$\frac{\lambda!}{x_0!x_1!x_2! \cdots x_k!} \quad (17)$$

b) Arranjo das n bolas por cada um dos k grupos de caixas com $s.x_s$ bolas:

$$\frac{n!}{x_1!(2x_2)! \cdots (kx_k)!} \quad (18)$$

c) Produto das k razões: permutações das $s.x_s$ bolas do grupo s pelas permutações das x_s posições das caixas do mesmo grupo ($s = 1, 2, \dots, k$):

$$\frac{x_1!}{1!} \frac{(2x_2)!}{(2!)^{x_2}} \frac{(3x_3)!}{(3!)^{x_3}} \cdots \frac{(kx_k)!}{(k!)^{x_k}} \quad (19)$$

Como no total são λ^n ocupações possíveis, então a probabilidade é o produto dos três fatores em relação ao total de ocupações possíveis. Simplificando o produto do segundo e do terceiro fatores, e reagrupando os termos, a probabilidade de uma ocupação é dada por:

$$P(x_0, x_1, x_2, \dots, x_k) = \frac{\lambda!}{x_0!x_1!x_2! \cdots x_k!} \frac{n!}{1!(2!)^{x_2}(3!)^{x_3} \cdots (k!)^{x_k}} \frac{1}{\lambda^n} \quad (20)$$

O primeiro fator está relacionado ao número de arranjos das λ caixas, vazias e com bolas; o segundo fator está relacionado ao número de arranjos das bolas pelos k grupos de bolas e por último a probabilidade de uma ocupação em relação total de ocupações possíveis.

Para o caso que todas as x caixas ocupadas tiverem uma única bola, ou seja, $x_1 = x \geq 1$ e $k = 1$ implica que há $\lambda - x$ caixas vazias e $n = x$ bolas, logo:

$$P(\lambda - n, n) = \frac{\lambda!}{(\lambda - n)!n!} \frac{n!}{1} \frac{1}{\lambda^n}$$

$$P(\lambda - n, n) = \frac{\lambda!}{(\lambda - n)! \lambda^n} = \frac{\lambda^{(n)}}{\lambda^n} = \frac{P_{K_n}(\lambda)}{\lambda^n} \quad (21)$$

Este resultado é o complemento do problema clássico dos aniversários, sendo que o numerador é o polinômio cromático de um grafo completo de n vértices. Como bolas e caixas são distintas, uma bola em cada caixa significa uma cor diferente para cada vértice, ou seja, uma coloração própria.

4.2 Problema clássico dos aniversários com múltiplos eventos

A discussão do trabalho a seguir ajuda a ampliar a questão da semelhança entre a solução combinatória e a por polinômios cromáticos, do ponto de vista das coincidências. Neste ponto, se for adotado que ocorre uma coincidência quando uma mesma cor é atribuída a dois vértices vizinhos, isto equivale a que pelo menos duas bolas ocupem a mesma caixa.

O artigo de Hocking e Schwertman (1986) tem como objetivo principal ser um exemplo didático das técnicas de contagem e propõe o caso onde y caixas têm duas bolas, todas as outras caixas ocupadas têm uma única bola, ou seja, $x_2 = y \geq 1$, $x_1 = n - 2y \geq 1$ e $k = 2$ implica que há $\lambda - n + y \geq 0$ caixas vazias e $x_1 + 2x_2 = n \geq 1$ bolas. Utilizando a solução proposta por von Mises (equação 20):

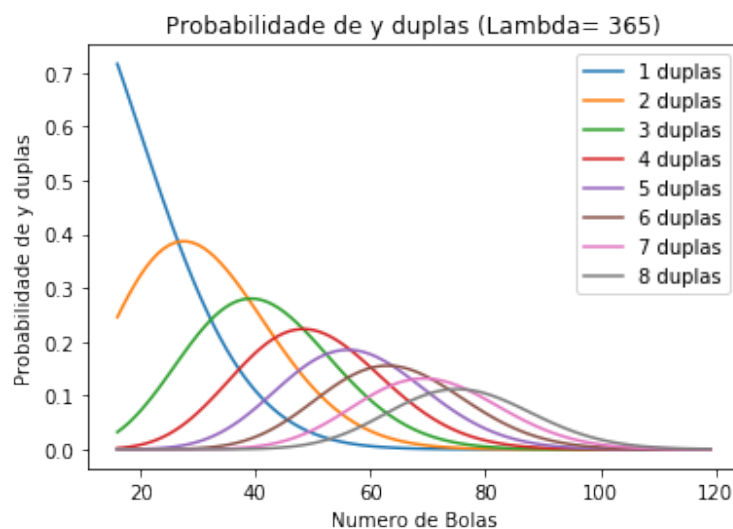
$$P(\lambda - n + y, n - 2y, y) = \frac{\lambda!}{(\lambda - n + y)!(n - 2y)! y!} \frac{n!}{2^y} \frac{1}{\lambda^n} \quad (22)$$

Rearranjando os termos de forma que o segundo fator agrupe os termos associados a λ e utilizando o símbolo de Pochhammer para $\lambda^{(n-y)} = \lambda! / (\lambda - n + y)!$ resulta:

$$P(\lambda - n + y, n - 2y, y) = \frac{\lambda!}{(\lambda - n + y)!} \frac{n!}{2^y (n - 2y)! y!} \frac{1}{\lambda^n} = \frac{n!}{(n - 2y)! 2^y y!} \frac{\lambda^{(n-y)}}{\lambda^n} \quad (23)$$

Hocking e Schwertman (1986) apresentam a probabilidade $P(\lambda - n + y, n - 2y, y)$ para valores de número de caixas com 2 bolas $x_2 = y$ entre 1 e 8 e o restante das bolas em caixas com uma bola, até o limite de preencher todas as λ caixas.

Figura 4.2.1 – Probabilidade de y coincidências duplas



Fonte: Elaboração própria

O anexo B apresenta o programa para o cálculo das curvas da Figura 4.2.1.

Para $y = 1$ tem-se:

$$P(\lambda - n + 1, n - 2, 1) = \frac{n!}{2(n-2)!} \frac{\lambda^{(n-1)}}{\lambda^n} = \frac{n(n-1)}{2} \frac{\lambda^{(n-1)}}{\lambda^n} \quad (24)$$

Imaginando um grafo completo com coloração própria, o primeiro fator sugere uma contagem relativa às $n(n-1)/2$ possibilidades de se obter uma coincidência simples, ou seja, dois vértices com a mesma cor. Esta coincidência pode ser obtida pela contração de uma aresta cujos vértices têm a mesma cor.

O segundo fator sugere a razão do número de colorações próprias do grafo completo resultante da contração de uma aresta (coincidência simples) sobre o número de possibilidades total de colorações. Vale lembrar que o polinômio cromático expressa o número de colorações próprias.

A tabela 1 mostra os casos de coincidências duplas e triplas de um grafo completo, utilizando a solução da equação de von Mises (equação 20):

Tabela 1 – Probabilidades, Coincidências e Contração de arestas

| $P(x_0, x_1, \dots, x_k)$ | Coincidências | Contrações de arestas | Caixas Ocupadas | Fator de contagem | Fator de Contração |
|--------------------------------------|----------------------|----------------------------|-----------------|---|-------------------------------------|
| $P(\lambda - n, n)$ | nenhuma | nenhuma | n | 1 | $\frac{\lambda^{(n)}}{\lambda^n}$ |
| $P(\lambda - n + 1, n - 2, 1)$ | 1 simples | 1 | n - 1 | $\frac{n!}{(n-2)! 2! 1!} = \frac{n^{(2)}}{2! 1!}$ | $\frac{\lambda^{(n-1)}}{\lambda^n}$ |
| $P(\lambda - n + 2, n - 3, 0, 1)$ | 1 dupla | 2 conexas | n - 2 | $\frac{n!}{(n-3)! 3! 1!} = \frac{n^{(3)}}{3! 1!}$ | $\frac{\lambda^{(n-2)}}{\lambda^n}$ |
| $P(\lambda - n + 2, n - 4, 2)$ | 2 simples | 2 desconexas | n - 2 | $\frac{n!}{(n-4)! (2!)^2 2!} = \frac{n^{(4)}}{(2!)^2 2!}$ | $\frac{\lambda^{(n-2)}}{\lambda^n}$ |
| $P(\lambda - n + 3, n - 4, 0, 0, 1)$ | 1 tripla | 3 conexas | n - 3 | $\frac{n!}{(n-4)! 4! 1!} = \frac{n^{(4)}}{4! 1!}$ | $\frac{\lambda^{(n-3)}}{\lambda^n}$ |
| $P(\lambda - n + 3, n - 5, 1, 1)$ | 1 dupla 1 simples | 2 conexas e 1 desconexa | n - 3 | $\frac{n!}{(n-5)! 3! 2!} = \frac{n^{(5)}}{3! 2!}$ | $\frac{\lambda^{(n-3)}}{\lambda^n}$ |
| $P(\lambda - n + 3, n - 6, 3)$ | 3 simples | 3 desconexas | n - 3 | $\frac{n!}{(n-6)! (2!)^3 3!} = \frac{n^{(6)}}{(2!)^3 3!}$ | $\frac{\lambda^{(n-3)}}{\lambda^n}$ |

Fonte: Elaboração própria

Na tabela 1 observa-se que o fator de contração está relacionado ao número de caixas ocupadas, que é o maior expoente do polinômio cromático. A redução do número de caixas ocupadas significa que uma bola que ocupava uma caixa x_1 e passa a ocupar uma outra caixa

já ocupada, por exemplo: no caso da redução de 3 casas ocupadas, estas bolas podem estar todas em x_4 (coincidência tripla), ou 2 em x_3 (coincidências duplas) e uma em x_2 (coincidência simples), ou ainda todas em x_2 (coincidências simples). Quando as bolas são reagrupadas em uma única caixa é quando a contração ocorre em um único caminho conexo, quando as bolas são reagrupadas em mais de uma caixa há tantos caminhos quantos foram as caixas que recebem as bolas.

É importante ressaltar que o problema clássico dos aniversários é representado pelo grafo completo, ou seja, cada caixa é vizinha de todas as outras, o que só é possível em um "espaço n -dimensional". De outra forma, uma bola pode ser transferida para qualquer outra caixa porque todas são suas vizinhas. No caso de um grafo qualquer, esta transferência só é possível através das arestas incidentes ao vértice a ser contraído.

4.3 Outro problema clássico dos aniversários

Mckinney (1966) propõe uma generalização do problema dos aniversários, onde as pessoas são adicionadas uma a uma até que k delas tenham a mesma data de aniversário.

A formulação proposta do problema é: qual é o menor número requerido de pessoas para que pelo menos k pessoas tenham a mesma data de aniversário com probabilidade $1/2$?

A solução proposta utiliza basicamente as mesmas hipóteses e argumentos da solução de von Mises (1939). Para um evento E : {não há mais do que uma caixa com k bolas ou mais que k bolas}, a probabilidade procurada é a do evento \bar{E} : {uma caixa tem k bolas e as outras têm menos bolas} = $1 - P(E)$. Portanto:

$$x_k = 1, \quad x_{k+1} = x_{k+2} = \dots = x_n = 0$$

A restrição do número de bolas $n = \sum_{i=1}^k i n_i$ pode ser atendida por todas as partições de n com parcelas menores que k , portanto o cálculo do número das possibilidades é a soma das ocupações das λ caixas por n bolas distribuídas conforme cada uma das partições de n com parcelas menores que k , ou seja:

$$x_1 + 2.x_2 + \dots + (k-1).x_{k-1} + 1.x_k = n$$

Por exemplo, para $n = 10$, $k = 3 \Rightarrow x_3 = 1 \Rightarrow 3$ bolas e as partições de $n - k = 7$ são:

| | |
|-------------------------|--------------------|
| $[2, 2, 2, 1]$ | $x_1 = 1, x_2 = 3$ |
| $[2, 2, 1, 1, 1]$ | $x_1 = 3, x_2 = 2$ |
| $[2, 1, 1, 1, 1, 1]$ | $x_1 = 5, x_2 = 1$ |
| $[1, 1, 1, 1, 1, 1, 1]$ | $x_1 = 7, x_2 = 0$ |

Mckinney (1966) apresenta uma tabela limitada $k = 2, 3$ e 4 . Não foram apresentados resultados para $k > 4$ devido ao excessivo tempo de processamento em um computador IBM 7090 em 1966.

O anexo C apresenta o programa para do número de pessoas para os mesmos caso e foi encontrada a mesma dificuldade com os recursos disponíveis na execução deste trabalho. A Tabela 2 permite avaliar o crescimento do número de parcelas das partições em relação número de bolas para cada incremento do número coincidências k . O valor do número de bolas para $k = 5$ também não foi calculado, principalmente após o cálculo do número de partições para $n = 365$.

Tabela 2 – Número de pessoas e Probabilidades para k delas com mesma data de aniversário

| k | 2 | | 3 | | 4 | | 5 |
|--------------------|--------|--------|--------|--------|--------|--------|--------|
| n | 22 | 23 | 87 | 88 | 186 | 187 | 365 |
| Probabilidade | 0,4757 | 0,5073 | 0,4995 | 0,5111 | 0,4958 | 0,5027 | * |
| Partições | 1 | 1 | 44 | 45 | 2977 | 3008 | 351726 |
| Tempo Proc. Aprox. | 0,01 s | | 90 s | | 1,6 Ms | | * |

*: não foi calculada

4.4 Problema dos quase aniversários

Abramson e Moser (1970) apresenta uma variação do problema dos aniversários: em um grupo de n pessoas, escolhidas aleatoriamente, qual a probabilidade de que cada par de pessoas tenham data de aniversário em um intervalo de, no mínimo, k dias?

A probabilidade do evento complementar é um problema de quase aniversário: em um grupo de n pessoas, escolhidas aleatoriamente, qual a probabilidade de que cada par de pessoas tenham data de aniversário em um intervalo menor que k dias?

As hipóteses são as clássicas acrescidas de que as datas do fim de um ano com o início do ano seguinte são consecutivas.

A solução resulta da análise das permutações dos $n - 1$ intervalos sujeitos as restrições $k \leq |b_j - b_i| \leq \lambda - k, 1 \leq i \neq j \leq n - k$, onde $b_i, i = 1, 2, \dots, n$ são as datas de aniversários das n pessoas e λ é número de dias do ano. A solução apresentada é dada por:

$$P(\lambda, n) = \frac{(n-1)!}{\lambda^{n-1}} \binom{\lambda - n(k-1) - 1}{n-1} \quad (25)$$

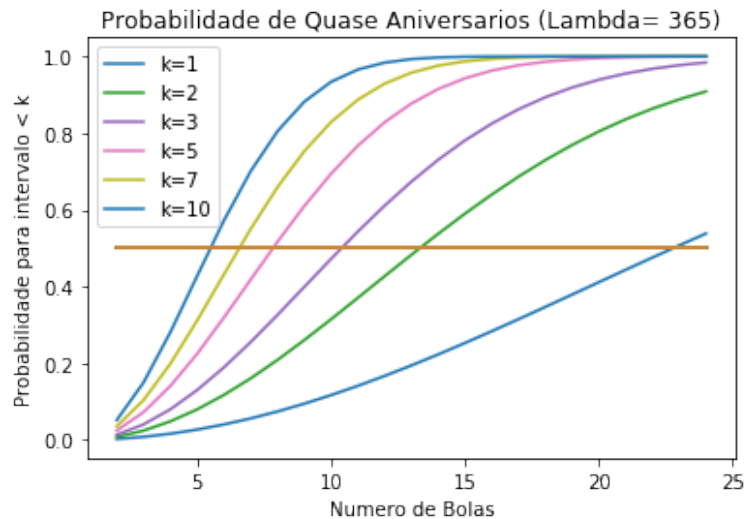
Esta equação pode ser reescrita utilizando a notação de Pochhammer:

$$P(\lambda, n) = \frac{(\lambda - n(k-1) - 1)^{(n-1)}}{\lambda^{n-1}} \quad (26)$$

É interessante verificar que aceitar um intervalo de k dias equivale a "encurtar" o ano de λ para $(\lambda - n(k-1) - 1)$, ou na linguagem das caixas e bolas, significa ter menos $n(k-1) + 1$ caixas disponíveis.

Um exemplo numérico para $\lambda = 365$ e k entre 1 e 10 é apresentado na figura 4.4.1. Este exemplo mostra que a probabilidade de quase aniversário aumenta rapidamente com o aumento do intervalo k . A tabela 3 mostra o número de pessoas para a probabilidade de quase aniversário $P(\lambda, n) \geq 1/2$ para intervalos menores que k dias.

Figura 4.4.1 – Probabilidade de quase aniversários com intervalo menor que k



Fonte: Elaboração própria

Tabela 3 – Número de pessoas para $P(\lambda, n) \geq 1/2$ com intervalo menor que k dias

| k | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---------------------------------|----|----|----|---|---|---|---|---|---|----|
| n para $P(\lambda, n) \geq 1/2$ | 23 | 14 | 11 | 9 | 8 | 8 | 7 | 7 | 6 | 6 |

4.5 Variações do problema dos aniversários

A literatura sobre o problema dos aniversários é vasta, muitos trabalhos utilizam conceitos matemáticos avançados, ou seja, de conhecimento de especialistas no assunto. Um exemplo desta diversidade de abordagens é apresentada por DasGupta (2005), que inclui a forma canônica, os problemas com distribuições não identicamente distribuídas, uma versão bayesiana desenvolvida por Diaconis e Holmes (2002) e as aproximações baseadas na distribuição de Poisson, como em Diaconis e Mosteller (1989). Neste trabalho somente a abordagem clássica ou canônica é tratada.

Seguem alguns pontos importantes dos resultados qualitativos destas abordagens.

O primeiro é a não uniformidade das distribuições de probabilidades em Rust (1976), Berresford (1980) e Nunnikhoven (1992). O primeiro inclui o efeito dos anos bissextos, que têm o impacto de aumentar a probabilidade de uma coincidência. O segundo mostra que as distribuições não-uniformes das datas de aniversários durante o ano têm maiores probabilidades que a uniforme, além disso apresenta um exemplo com dados dos nascimentos em Nova York no ano de 1977. Na mesma linha, Nunnikhoven (1992) propõe um modelo composto pela solução clássica à qual são adicionados uma série de termos que rapidamente se aproximam de zero.

Fadnavis (2015) também demonstra que as distribuições uniformes provocam um aumento da probabilidade de uma coincidência e é o único trabalho identificado na literatura que associa o problema dos aniversários à coloração de grafos, porém seu foco é na demonstração dos efeitos da não uniformidade da distribuição em grafos simétricos, como os completos e as estrelas.

Outro grupo de trabalhos, que não fazem parte do escopo deste trabalho, são os que tratam da solução do problema dos aniversários com a utilização de aproximações. Diaconis e Mosteller (1989) apresenta aproximações para os problemas clássicos, de muitas categorias, de múltiplos exemplos e de quase-aniversários. Este tema continua sendo estudado, como pode ser visto em, por exemplo, Bhattacharya, Diaconis e Mukherjee (2017).

A atual situação dos algoritmos, dos sistemas de cálculos e da capacidade computacional, tanto de memória quanto de velocidade de processamento, restringem a solução de problemas dos aniversários a casos simples, o que torna interessante o estudo das aproximações e de métodos que utilizam números aleatórios. Estas possibilidades são extensões possíveis para este trabalho.

5 ALGORITMOS PARA CÁLCULO DE POLINÔMIOS CROMÁTICOS

O teorema fundamental e seu corolário definem duas formas para calcular o polinômio cromático de um grafo simples G : pela adição de arestas e pela remoção de arestas.

Dado um número de vértices, um grafo completo tem o número máximo de arestas para um grafo simples, enquanto as árvores têm o número mínimo.

Se for considerado o número de arestas como um parâmetro básico, é razoável concluir que para grafos densos, ou seja, com um número de arestas próximo ao número máximo de arestas $n(n-1)/2$ de um grafo completo com n vértices, a adição de arestas com base no teorema fundamental tem o potencial de ser mais efetiva.

O reverso ocorre para grafos esparsos, ou seja, com um número de arestas próximo ao mínimo de arestas $(n-1)$ de uma árvore conexa com n vértices, a remoção de arestas com base no corolário fundamental tem o potencial de ser mais efetiva.

Esta análise identifica três características básicas para caracterizar um grafo do ponto de vista do seu polinômio cromático:

- (a) número de vértices n ,
- (b) número de arestas m ,
- (c) densidade do grafo: a razão do número de arestas em relação ao número máximo de arestas de um grafo simples $n(n-1)/2$.

Porém, uma questão central para a aplicação do teorema ou do corolário fundamental é escolher a sequência de adição/remoção das arestas.

Este problema fundamental é estudado e revisado na literatura: (NIJENHUIS; WILF, 1978), (READ, 1987), (CHIA, 1997), (HAGGARD; PEARCE; ROYLE, 2010), (MONAGAN, 2012), (LEWIS, 2015), (LIMA, 2017), entre muitas outras referências.

Uma característica importante dos polinômios cromáticos é que seus coeficientes são números inteiros e o sinal de coeficientes sucessivos têm sinais contrários, o que pode ser associado à sua propriedade adição/contração ou remoção/contração. Portanto estes coeficientes podem alcançar centenas de dígitos. Por exemplo, o polinômio cromático de um grafo completo de $n=30$ vértices é dado na figura 5.0.1.

A seleção do sistema de cálculo deve ser cuidadosa de forma que os erros de arredondamento não comprometam os cálculos, além de outra dificuldade relatada em Nijenhuis e Wilf (1978) que é "overflow" durante cálculos que envolvem fatoriais e multiplicações de números

Figura 5.0.1 – Coeficientes na base de potência do polinômio cromático completo de $n = 30$ vértices

$$\begin{aligned}
& +\lambda^{30} - 435 * \lambda^{29} \\
& +90335 * \lambda^{28} - 11921175 * \lambda^{27} \\
& +1122686019 * \lambda^{26} - 80328850875 * \lambda^{25} \\
& +4539323721075 * \lambda^{24} - 207912996295875 * \lambda^{23} \\
& +7860403394108265 * \lambda^{22} - 248526574856284725 * \lambda^{21} \\
& +6634460278534540725 * \lambda^{20} - 150566737512021319125 * \lambda^{19} \\
& +2918939500751087661105 * \lambda^{18} - 48487623689430693038025 * \lambda^{17} \\
& +691254538651580660999025 * \lambda^{16} - 8459574446076318147830625 * \lambda^{15} \\
& +88776380550648116217781890 * \lambda^{14} - 796974693974455191377937300 * \lambda^{13} \\
& +6097272817323042122728617800 * \lambda^{12} - 39539238727270799376544542000 * \lambda^{11} \\
& +215760462268683520394805979744 * \lambda^{10} - 981347603630155088295475765440 * \lambda^9 \\
& +3674201658710345201899117607040 * \lambda^8 - 11139316913434780466101123891200 * \lambda^7 \\
& +26751280755793398822580822142976 * \lambda^6 - 49361465831621147825759587123200 * \lambda^5 \\
& +66951000306085302338993639424000 * \lambda^4 - 62262192842035613491057459200000 * \lambda^3 \\
& +35027999979859805266492784640000 * \lambda^2 - 8841761993739701954543616000000 * \lambda
\end{aligned}$$

Fonte: Elaboração própria

muito grandes. O uso de números reais como coeficientes minimiza os problemas de "overflow", porém exige uma análise cuidadosa dos erros de arredondamento. A escolha de permanecer com polinômios com coeficientes de números inteiros limita as opções aplicáveis no estudo de polinômios cromáticos, preserva a exatidão dos cálculos, entretanto limita a dimensão das redes devido ao longo tempo de processamento e eventual uso de memória.

No decorrer do trabalho foram exploradas possibilidades de implementar soluções conhecidas que permitissem que o esforço se concentrasse no tema principal e, secundariamente, nas simulações. Como a parte do aprendizado de testar limites e novas ideias é essencial, foram avaliados prioritariamente ambientes como: Phyton, R, SageMath, Mathematica, Maple, Magma, MatLab, entre outros, ou seja, as linguagens como C, Java, Julia não foram consideradas opções viáveis para este trabalho, uma vez que o autor não é especialista em ciência da computação e nem em programação.

Outro critério considerado relevante foi a preferência por ambientes de software livre e de código aberto, assim as opções se reduziram a: Phyton, R, SageMath.

Outro critério, não identificado inicialmente, se mostrou muito importante, que é a capacidade de realizar processamento simbólico, notadamente de polinômios de coeficientes inteiros. Este critério elimina todas as linguagens tradicionais. A contrapartida de ter disponível o cálculo simbólico e relevantes bibliotecas de funções específicas a grafos e análise combinatória é a limitação da dimensão das redes devido aos longos tempos de processamento.

SageMath é um ambiente escrito em Python, porém com as bibliotecas e pacotes compilados, o que permite uma maior velocidade nos cálculos. Além disso, podem ser utilizadas todas as funções do Python, do R, e incorporar módulos escritos em outras linguagens.

O SageMath é capaz de operar cálculos simbólicos e com qualquer número de dígitos inteiros, limitado pela memória do computador utilizado. Outro ponto importante é a biblioteca de funções e algoritmos relacionados a grafos em geral e polinômios cromáticos, em particular.

Por ser de código aberto, sabe-se que a rotina de cálculo de polinômios cromáticos é baseada no relatório Read (1987) e foi implementada em 2008 por Robert Miller e Gordon Royle. Utiliza o método da remoção/contração com um teste se o grafo contraído é uma árvore, simplesmente verificando se o número de arestas é o número de vértices menos um. Miller e Gordon relataram que, esta simples verificação e a solução analítica dos polinômios cromáticos de árvores trouxeram melhoria de desempenho nos casos de grafos esparsos.

Durante o estudo dos teoremas e as tentativas de modelagem dos problemas estudados foi verificado que grafos esparsos, porém com pontos de concentração, como "hubs", podem ser resolvidos mais rapidamente com abordagens híbridas, ou seja, com redução das partes da rede que são esparsas, porém a partir de uma certa densidade, adoção do método de adição/contração.

Estudar outras heurísticas relacionadas a polinômios cromáticos, como: a proposta por Monagan (2012), remoção de arestas para gerar árvores, triangularização e verificação de grafos cordais (tem uma sequência linear de remoção/contração) são extensões possíveis deste trabalho.

A utilização do SageMath (versão 8.8) foi baseada em notebook Jupyter, rodando em Linux Mint em computador pessoal.

A percepção do autor é que seu aprendizado foi qualitativamente mais profundo e mais rápido a partir da utilização do ambiente Sagemath. É razoável imaginar que este trabalho não teria sido possível dentro do prazo disponível, ou teria um escopo mais reduzido.

6 CASOS DE APLICAÇÃO

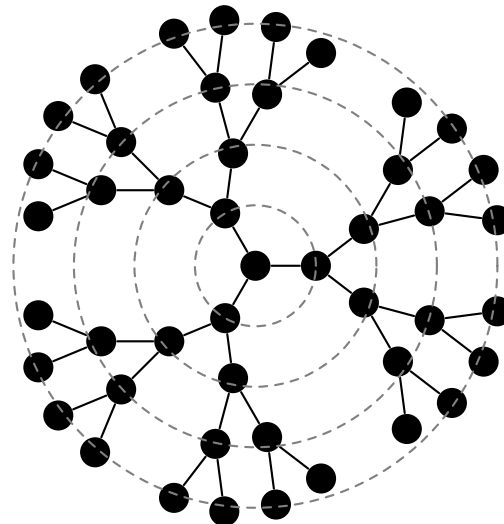
6.1 Grafos Árvores de Cayley

Baxter (1982) relaciona as árvores de Cayley e rede de Bethe ao modelo de Ising nos estudos da mecânica estatística e propõe uma definição de uma árvore de Cayley a partir da sua construção:

- comece de um ponto central, denominado 0,
- adicione q vértices e os conecte com o ponto central,
- denomine o conjunto dos vértices adicionados de 1^a camada ou geração,
- para cada vértice da camada anterior, adicione $(q - 1)$ vértices e os conecte com este vértice,
- denomine o conjunto dos vértices adicionados de 2^a camada ou geração,
- repita os passos d) e e) quantas vezes forem necessárias.

Um exemplo da primeira camada com $q = 3$ vértices e as demais com $q - 1 = 2$ é apresentado a seguir. Nesta construção há $q(q - 1)^{r-1}$ vértices na camada r .

Figura 6.1.1 – Árvore de Cayley de ordem 2: $q=3$, $q-1= 2$ com 4 camadas



Fonte: Elaboração própria

O polinômio cromático de qualquer árvore é dado por $P(\lambda) = \lambda(\lambda - 1)^{n-1}$. Fixada a cor de um vértice qualquer, as possibilidades de coloração dos seus vizinhos diminui de uma cor. Como não há circuitos em uma árvore, então este argumento vale para todos os vértices sucessivamente, o vizinho do vizinho recupera a possibilidade de coloração do vértice inicial.

Um exemplo genérico permite esclarecer um dos argumentos centrais deste trabalho: a importância da transformação de base do polinômio. Seja uma árvore qualquer com 50 vértices.

(a) na base árvore:

$$P(\lambda) = \lambda(\lambda - 1)^{49}$$

(b) na base fatorial:

$1 \lambda^{(50)}$
 $+ 1176 \lambda^{(49)}$
 $+ 654052 \lambda^{(48)}$
 $+ 229037956 \lambda^{(47)}$
 $+ 56723760534 \lambda^{(46)}$
 $+ 10580260064220 \lambda^{(45)}$
 $+ 1545513894055670 \lambda^{(44)}$
 $+ 181521707399014290 \lambda^{(43)}$
 $+ 17465076236169935065 \lambda^{(42)}$
 $+ 1395530153844068098780 \lambda^{(41)}$
 $+ 9355674288874819048950 \lambda^{(40)}$
 $+ 5303066770830294372037950 \lambda^{(39)}$
 $+ 255632810197624387599902145 \lambda^{(38)}$
 $+ 10524796006195412206328516670 \lambda^{(37)}$
 $+ 371243032069067047342741410675 \lambda^{(36)}$
 $+ 11242030196128777844505818320125 \lambda^{(35)}$
 $+ 292604389757333075769183556660710 \lambda^{(34)}$
 $+ 6548306522892167432294799248101680 \lambda^{(33)}$
 $+ 125960270431065386748055072930933760 \lambda^{(32)}$
 $+ 2080355002188598095900300532972416000 \lambda^{(31)}$
 $+ 29450293801972699624155238989695481752 \lambda^{(30)}$
 $+ 356488166397691879910086909729794416912 \lambda^{(29)}$
 $+ 3678386002142588376130895185580929170264 \lambda^{(28)}$
 $+ 32230205307958519658280688245780281858952 \lambda^{(27)}$
 $+ 238704263590751290535414759819485245995460 \lambda^{(26)}$
 $+ 1486195563504417819994257239350494316026000 \lambda^{(25)}$
 $+ 7729050276141867153694397233020421577807400 \lambda^{(24)}$
 $+ 33324337798684426660637559740175410878576200 \lambda^{(23)}$
 $+ 118086451506299247707595962362400131422949980 \lambda^{(22)}$
 $+ 340439546931221947798437745027410288290773800 \lambda^{(21)}$
 $+ 789130668236547944485714493551071381183279300 \lambda^{(20)}$
 $+ 145051095543066340728382684340593776734790300 \lambda^{(19)}$
 $+ 2080135615741512810791093553751110341051310765 \lambda^{(18)}$
 $+ 2282936323995189021422490404856161928375236040 \lambda^{(17)}$
 $+ 1873844181572520481087961510837480961776274300 \lambda^{(16)}$
 $+ 111879427837331873474430550564295843380185500 \lambda^{(15)}$
 $+ 469689372783848147483295280135550173835893930 \lambda^{(14)}$
 $+ 132939430085758894210023153064585304022667140 \lambda^{(13)}$
 $+ 24057783907173048634103909041404388169280730 \lambda^{(12)}$
 $+ 260013111118083676032277591483101188106750 \lambda^{(11)}$
 $+ 153405404682044531690001481746981900659511 \lambda^{(10)}$
 $+ 437386765755399391893753977893403744596 \lambda^{(9)}$
 $+ 50789872166903636182659702516635946082 \lambda^{(8)}$
 $+ 186954124427011518447120396686063706 \lambda^{(7)}$
 $+ 148016531942876772055857596434451 \lambda^{(6)}$
 $+ 13204653869155991900157415370 \lambda^{(5)}$
 $+ 39883221256961278221025 \lambda^{(4)}$
 $+ 281474976710655 \lambda^{(3)}$
 $+ 1 \lambda^{(2)}$
 $+ 0 \lambda^{(1)}$

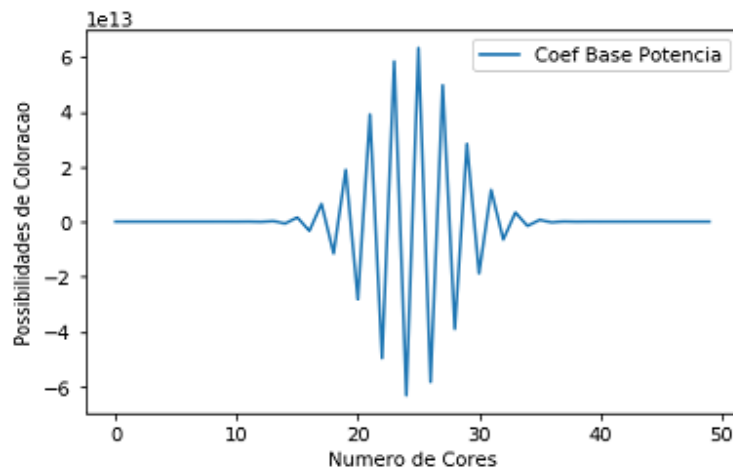
(c) na base de potência:

$\lambda^{50} - 49\lambda^{49} + 1176\lambda^{48} - 18424\lambda^{47} + 211876\lambda^{46} - 1906884\lambda^{45} + 13983816\lambda^{44} - 85900584\lambda^{43} + 450978066\lambda^{42} +$
 $- 2054455634\lambda^{41} + 8217822536\lambda^{40} - 29135916264\lambda^{39} + 92263734836\lambda^{38} - 262596783764\lambda^{37} + 675248872536\lambda^{36} +$
 $- 1575580702584\lambda^{35} + 3348108992991\lambda^{34} - 6499270398159\lambda^{33} + 11554258485616\lambda^{32} - 18851684897584\lambda^{31} +$
 $+ 28277527346376\lambda^{30} - 39049918716424\lambda^{29} + 49699896548176\lambda^{28} - 58343356817424\lambda^{27} + 63205303218876\lambda^{26} +$
 $- 63205303218876\lambda^{25} + 58343356817424\lambda^{24} - 49699896548176\lambda^{23} + 39049918716424\lambda^{22} - 28277527346376\lambda^{21} +$
 $+ 18851684897584\lambda^{20} - 11554258485616\lambda^{19} + 6499270398159\lambda^{18} - 3348108992991\lambda^{17} + 1575580702584\lambda^{16} +$
 $- 675248872536\lambda^{15} + 262596783764\lambda^{14} - 92263734836\lambda^{13} + 29135916264\lambda^{12} - 8217822536\lambda^{11} +$
 $+ 2054455634\lambda^{10} - 450978066\lambda^9 + 85900584\lambda^8 - 13983816\lambda^7 + 1906884\lambda^6 - 211876\lambda^5 +$
 $+ 18424\lambda^4 - 1176\lambda^3 + 49\lambda^2 - \lambda$

A análise dos coeficientes na base fatorial mostra que cada termo $\lambda^{(k)}$ representa o número de colorações próprias com exatas k cores, porém as colorações são indiferentes, ou seja, não há distinção nas rotações das cores. Para obter o número de colorações próprias com diferenciação das cores basta multiplicar o coeficiente por $k!$. O menor expoente do fatorial decrescente é o número cromático, no caso: $\chi(G) = 2$.

Os coeficientes da base de potência são valores discretos, se colocados em um gráfico com os seus valores conectados por linhas fica clara a alternância dos sinais e a variação relativa entre eles, como pode ser visto em 6.1.2.

Figura 6.1.2 – Variação dos coeficientes da base de potência



Fonte: Elaboração própria

O teorema 3.5.2 permite interpretar os coeficientes da base de potência como colorações altamente impróprias, ou seja, são as coincidências que se distribuem como florestas de circuitos quebrados, onde para cada expoente as florestas têm o mesmo número de componentes e também o mesmo total de arestas.

Como foi visto no caso dos problemas dos aniversários com múltiplos eventos, seção 4.2, a conexidade dos componentes da floresta altera sua contribuição na contagem dos casos possíveis, portanto uma outra abordagem é necessária para estudar os subgrafos de G , que não contenham nenhum circuito quebrado.

Até este ponto as afirmações são válidas para qualquer árvore, porém para estudar os subgrafos é necessário considerar a topologia de G . A seguir é apresentado um exemplo de uma abordagem específica para as árvores de Cayley.

Pah (2008) estuda o número de componentes conexos de acordo com o número de vértices na árvore de Cayley semi-infinita de ordem 2, que se inicia em um vértice x_0 da primeira camada e segue por "infinitas" camadas.

A distância $d(u, v)$ entre dois vértices u e v da árvore G é definida como o número de arestas contadas no caminho entre u e v , que existe devido a conexidade da árvore e é único em uma árvore.

Seja $N(m)$ o número de subgrafos H induzidos conexos de G com m vértices que contém o vértice raiz x_0 .

Seja a numeração dos vértices de $V(G)$, onde o vértice raiz x_0 como 1, seus sucessores na árvore semi-infinita serão $2k$ e $2k + 1$, sucessivamente em cada camada, assim os vértices dos subgrafos que contém x_0 são representados pelas seguintes sequências:

$$\begin{aligned}
 N(m = 1) &= 1 && (1) \\
 N(m = 2) &= 2 && (1, 2), (1, 2) \\
 N(m = 3) &= 5 && (1, 2, 3), (1, 2, 4), (1, 2, 5), (1, 3, 6), (1, 3, 7) \\
 N(m = 4) &= 14 && (1, 2, 3, 4), (1, 2, 3, 5), (1, 2, 3, 6), (1, 2, 3, 7), (1, 2, 4, 5), \\
 &&& (1, 2, 4, 8), (1, 2, 4, 9), (1, 2, 5, 10), (1, 2, 5, 11), (1, 3, 6, 7), \\
 &&& (1, 3, 6, 12), (1, 3, 6, 13), (1, 3, 7, 14), (1, 3, 7, 15) \\
 \dots &&& \dots
 \end{aligned}$$

Os valores dos números de subgrafos $N(m)$ sugere que esta é uma sequência de Catalan. Stanley (2015) apresenta 214 exemplos de problemas que envolvem estes números, sendo que o número de árvores binárias com vértice raiz e $n + 1$ folhas é um número de Catalan C_n .

Pah (2008) apresenta o seguinte teorema.

Teorema 6.1.1 *O número de subgrafos H induzidos conexos de uma árvore de Cayley com m vértices, que contém o vértice raiz x_0 , $N(m)$ é dada pela seguinte relação de recorrência: $N(m) = \sum_{r=0}^m N(m-r-1)N(r)$, com $N(0) = 1$ e $N(1) = 1$.*

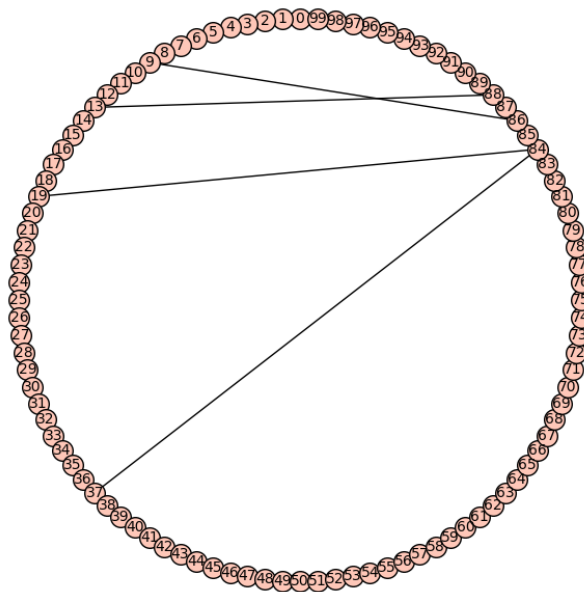
A interpretação é $N(r)$ e $N(m-r-1)$ são os sucessores de cada ramo da árvore a partir do vértice raiz, com números de vértices que somam m , logo o total de possibilidades é a soma destes produtos.

Utilizar a contagem apresentada para construir as florestas dos subgrafos de G , que não contenham nenhum circuito quebrado, é uma extensão possível para deste trabalho.

6.2 Grafos de Bollobas-Chung

Bollobas e Chung (1988) estudaram as características do diâmetro de grafos circulares com adição de arestas aleatórias, porém mantendo o grafo simples, ou seja, sem laços e sem arestas múltiplas. O diâmetro de uma rede está relacionado com a velocidade de troca de informações entre seus pontos (vértices) bem com sua fragilidade que pode ser interpretada como a impossibilidade de dois pontos se comunicar (conexidade) em caso de falha (perda de um vértice ou uma aresta da rede).

Figura 6.2.1 – Grafos de Bollobas-Chung



Fonte: Elaboração própria

Este modelo reúne três características que podem ser úteis na modelagem de redes reais simples (WANG; CHEN, 2003), (ESTRADA; KNIGHT, 2015):

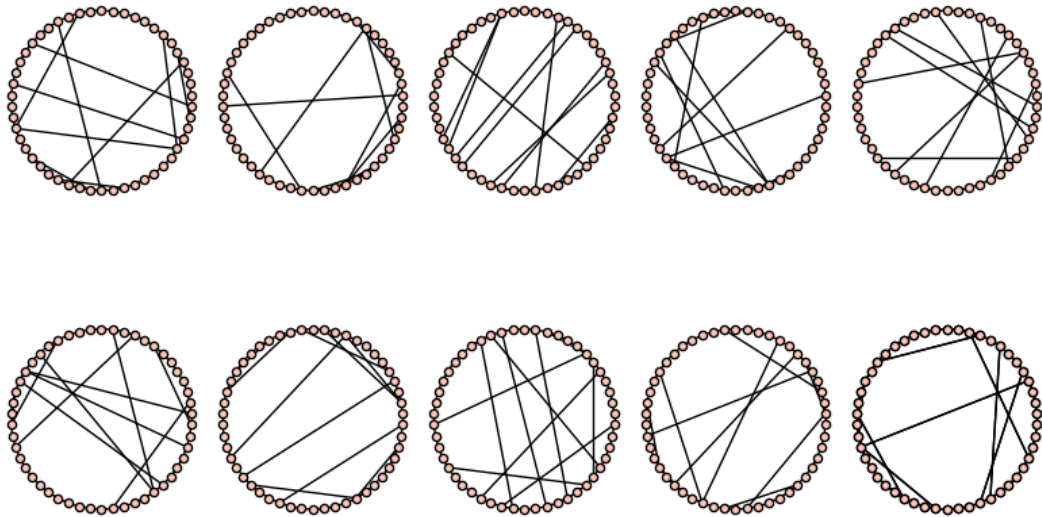
- (a) caminho médio entre vértices muito menor que o número de arestas (caminho mais longo possível), o que permite simular o efeito "small world" na rede,
- (b) distribuição de graus dos vértices que segue uma lei de potência, ou seja, rede sem escala ("scale-free"),
- (c) clusterização de vértices, ou seja, conectividade concentrada por grupos de vizinhos.

Estes três aspectos identificam características básicas de um grafo do ponto de vista do seu polinômio cromático: caminho médio entre vértices, a distribuição de graus e índice de clusterização.

Como ponto de partida foram gerados 10 grafos circulares com: $n= 20, 30, 40$ e 50 vértices. Em cada caso uma parte das arestas foram adicionadas aleatoriamente de forma a atingir a proporção de aresta aleatórias $p= 0,05; 0,10; 0,15$ e $0,20$.

Um exemplo dos grafos gerados é apresentado na Figura 6.2.2.

Figura 6.2.2 – Exemplos de grafo circular com 50 vértices $p= 0,20$ ou 10 arestas aleatórias



Fonte: Elaboração própria

Os coeficientes dos polinômios cromáticos podem ser vistos nas figuras 6.2.3 e 6.2.4.

A variação dos coeficientes dos polinômios na base de potência está relacionada aos circuitos quebrados e os coeficientes na base fatorial representam as colorações próprias com cores exatas.

A variação dos coeficientes, tanto nos grafos com 10 vértices como no de 20 vértices adicionados, é bastante limitada, contrariando a expectativa inicial que fosse relevante fazer uma análise destas variações. A limitação do tempo e outras oportunidades de aprendizado tornaram estas análise em um desdobramento futuro.

Uma possibilidade de desdobramento deste trabalho é: calcular as características já identificadas das redes (grafos), tais como: número de vértices, número de arestas, densidade, caminho médio, coeficiente de clusterização, distribuição dos graus e correlacionar com as características dos polinômios cromáticos, como: probabilidade de $k=0$ coincidências (total de rotações das colorações próprias), coeficientes da forma fatorial (número de rotações das colorações próprias para cada grafo completo), número cromático, e outras características a serem identificadas.

Figura 6.2.4 – Coeficientes dos polinômios cromáticos dos grafos circulares 50 vértices p=0,20

$x^{50} - 60x^{49} + 1770x^{48} - 34220x^{47} + 487635x^{46} - 5461512x^{45} + 50063855x^{44} - 386206649x^{43} + 2558613635x^{42} - 14783017159x^{41} + 75392420284x^{40} - 342683976995x^{39} + 1399226318644x^{38} - 5165950014593x^{37} + 17340503735960x^{36} - 53166362778634x^{35} + 149482893705419x^{34} - 386716787577473x^{33} + 923209362898975x^{32} - 203884679734730x^{31} + 4174402789585826x^{30} - 7937645498414067x^{29} + 14039375327103294x^{28} - 23126677253494004x^{27} + 35516848494300931x^{26} - 50893142247998057x^{25} + 68083350104646334x^{24} - 85062750873474496x^{23} + 99270743436955806x^{22} - 108207350262374321x^{21} + 110132334314916786x^{20} - 104605637168399666x^{19} + 9264373156998900x^{18} - 76418678014459505x^{17} + 58620373731009953x^{16} - 41736822794575086x^{15} + 27513312939546699x^{14} - 16740728385743096x^{13} + 9365385499469209x^{12} - 4793730210609067x^{11} + 2231172026873781x^{10} - 936865685920486x^9 + 351296657793301x^8 - 116059251628194x^7 + 33172826935281x^6 - 7996676566089x^5 + 1564457724851x^4 - 233660463328x^3 + 2375755338x^2 - 1237769980x$

$x^{50} - 60x^{49} + 1769x^{48} - 34162x^{47} + 485979x^{46} - 5430487x^{45} + 49634918x^{44} - 381540321x^{43} + 2517021675x^{42} - 14470242484x^{41} + 73377295092x^{40} - 331336000831x^{39} + 1342792496695x^{38} - 4915757722257x^{37} + 16344052773846x^{36} - 49578560490883x^{35} + 137742774210562x^{34} - 35164968809484x^{33} + 827242314379957x^{32} - 1797503987600842x^{31} + 3614909356743473x^{30} - 673987232727358x^{29} + 1166657796381308x^{28} - 18770205803819613x^{27} + 28094277314802336x^{26} - 39145195031777198x^{25} + 50797425386989643x^{24} - 6140478934327812x^{23} + 69144542170852532x^{22} - 72511500534087199x^{21} + 70785249654812137x^{20} - 64275691577643489x^{19} + 542346646571823259x^{18} - 42467212627930877x^{17} + 30806368264990832x^{16} - 20659420097258757x^{15} + 12774732073850594x^{14} - 726017262097131x^{13} + 3777379081089729x^{12} - 1790489206754089x^{11} + 768548060361315x^{10} - 296485099283652x^9 + 101807548263916x^8 - 3072879690612x^7 + 8016078014243x^6 - 1764892624270x^5 + 31651233771x^4 - 43620516496x^3 + 4136485292x^2 - 204041520x$

$x^{50} - 60x^{49} + 1769x^{48} - 34160x^{47} + 485867x^{46} - 5427410x^{45} + 49579642x^{44} - 380810144x^{43} + 2509459420x^{42} - 14406806566x^{41} + 72923622840x^{40} - 328579329575x^{39} + 1328228296299x^{38} - 4848055571815x^{37} + 16064513697607x^{36} - 48545465894522x^{35} + 134303766239179x^{34} - 341284155388220x^{33} + 798829892551873x^{32} - 1726420987933564x^{31} + 3452097380999152x^{30} - 6397602087578320x^{29} + 11004791101052178x^{28} - 17591345066649089x^{27} + 26157085996877364x^{26} - 36205729183129966x^{25} + 46676200145193823x^{24} - 56064323282404277x^{23} + 62748480785794199x^{22} - 65434385912520497x^{21} + 63556433056923059x^{20} - 57467909699848183x^{19} + 48334052820373630x^{18} - 37717169524367868x^{17} + 27387327254582088x^{16} - 1839147206892672x^{15} + 11412686598547712x^{14} - 652598382130852x^{13} + 3426719191552594x^{12} - 1645126996930344x^{11} + 718183866716829x^{10} - 283113056674695x^9 + 9987026826932x^8 - 31147053557087x^7 + 8446486488310x^6 - 1944615362565x^5 + 366533457231x^4 - 53270024801x^3 + 5330235980x^2 - 276550716x$

$x^{50} - 60x^{49} + 1770x^{48} - 34219x^{47} + 487576x^{46} - 5459803x^{45} + 50031464x^{44} - 385754931x^{43} + 2553671496x^{42} - 14738849074x^{41} + 75060913873x^{40} - 340551794311x^{39} + 1387293469819x^{38} - 5107141807146x^{37} + 17082839309163x^{36} - 52154847040990x^{35} + 145902064219110x^{34} - 375224186638393x^{33} + 889614710750644x^{32} - 1949119556730098x^{31} + 3954396375885729x^{30} - 744171351133335x^{29} + 13008785958515219x^{28} - 21148414304698216x^{27} + 32003402714741908x^{26} - 45111948500004137x^{25} + 59260618577401593x^{24} - 72564807418668769x^{23} + 8282854193575760x^{22} - 88112297253894170x^{21} + 87316849326147915x^{20} - 80547656576271834x^{19} + 69097733295100759x^{18} - 55049407843176072x^{17} + 4066123600330749x^{16} - 27785721142510863x^{15} + 1751977784694777x^{14} - 10159840345089636x^{13} + 5397014551346442x^{12} - 2613203166041988x^{11} + 1146198182859995x^{10} - 451884198830192x^9 + 158539083159012x^8 - 48854196798803x^7 + 12991166991977x^6 - 2907907197241x^5 + 527979456755x^4 - 7322299597x^3 + 6930513497x^2 - 337854950x$

$x^{50} - 60x^{49} + 1770x^{48} - 34219x^{47} + 487577x^{46} - 5459860x^{45} + 50033059x^{44} - 385784135x^{43} + 2554064964x^{42} - 14743008361x^{41} + 75096832443x^{40} - 340812322763x^{39} + 1388912981998x^{38} - 5115902059878x^{37} + 17124574139789x^{36} - 5233161483084x^{35} + 146522908523098x^{34} - 377519966253956x^{33} + 89674017813924x^{32} - 1969246044906174x^{31} + 4006383093400286x^{30} - 7564866076188759x^{29} + 13277081334935979x^{28} - 21687193798945439x^{27} + 3300268398251700x^{26} - 46826463853822464x^{25} + 61985504462084723x^{24} - 76580410397706244x^{23} + 88319775676899765x^{22} - 9508343674079585x^{21} + 95534058802885479x^{20} - 89539916999404822x^{19} + 782289881818631244x^{18} - 6364630904452894x^{17} + 48155954960541646x^{16} - 33825415666031412x^{15} + 22008487730416808x^{14} - 13227345262069473x^{13} + 7317284564308501x^{12} - 3709095899660319x^{11} + 171290130556421x^{10} - 71538999057130x^9 + 267567871061912x^8 - 88479507908199x^7 + 25413084211044x^6 - 6182271170121x^5 + 1226701451609x^4 - 186726830313x^3 + 19447019922x^2 - 1042886654x$

$x^{50} - 60x^{49} + 1770x^{48} - 34219x^{47} + 487577x^{46} - 5459859x^{45} + 50033004x^{44} - 385782650x^{43} + 2554038728x^{42} - 14742667251x^{41} + 75093352204x^{40} - 340783307217x^{39} + 1388707922505x^{38} - 5114681619716x^{37} + 17118188649206x^{36} - 52302181542366x^{35} + 146452181023275x^{34} - 377075854824338x^{33} + 8952653986671289x^{32} - 1964799065205732x^{31} + 3994157885119402x^{30} - 7534099837301525x^{29} + 13205969807187941x^{28} - 2153521605499635x^{27} + 32705254325406940x^{26} - 46286003803776544x^{25} + 6107595973040830x^{24} - 75161203511671952x^{23} + 86265049020361019x^{22} - 923230941028485x^{21} + 92089205153807826x^{20} - 85551848924679976x^{19} + 73947482214199350x^{18} - 59388406979440854x^{17} + 44239527018548014x^{16} - 30500396397943799x^{15} + 19409684019307957x^{14} - 11363536190445359x^{13} + 6095772078145909x^{12} - 2981240406898266x^{11} + 1321097633376572x^{10} - 526359008406899x^9 + 186706374579132x^8 - 58208550636142x^7 + 15676577854874x^6 - 3559406309593x^5 + 656960300701x^4 - 92867583318x^3 + 8985937434x^2 - 448929972x$

$x^{50} - 60x^{49} + 1770x^{48} - 34219x^{47} + 487577x^{46} - 5459859x^{45} + 50033002x^{44} - 385782543x^{43} + 2554035921x^{42} - 14742619122x^{41} + 75092745630x^{40} - 340777315556x^{39} + 1388661420950x^{38} - 5114354893632x^{37} + 17116296386142x^{36} - 52292653778600x^{35} + 14640997358799x^{34} - 3769097796812811x^{33} + 89468017909602x^{32} - 1962943453652163x^{31} + 3988827215202810x^{30} - 7520172530134876x^{29} + 13172755646357634x^{28} - 21463300564991440x^{27} + 32559918057425220x^{26} - 46018126090241613x^{25} + 60621123663158507x^{24} - 74448896244152720x^{23} + 85235234988855648x^{22} - 90946878831950086x^{21} + 90392567921892817x^{20} - 83618205789357065x^{19} + 71912611456258305x^{18} - 57412688805437185x^{17} + 42471737168594217x^{16} - 29045067597297150x^{15} + 18309625984061562x^{14} - 10602074671107652x^{13} + 5614696196276593x^{12} - 2704997158660434x^{11} + 1177690609337389x^{10} - 459513224549140x^9 + 158979797594746x^8 - 48097899729820x^7 + 12489295363360x^6 - 2711640359264x^5 + 47361226000x^4 - 62529538240x^3 + 5560765440x^2 - 250475520x$

$x^{50} - 60x^{49} + 1770x^{48} - 34220x^{47} + 487634x^{46} - 5461456x^{45} + 50062318x^{44} - 386179091x^{43} + 2558250635x^{42} - 1477921595x^{41} + 75360896418x^{40} - 342461475678x^{39} + 1397882409471x^{38} - 5158896534590x^{37} + 170073074273659x^{36} - 53032939455510x^{35} + 148993571712135x^{34} - 38510040501019x^{33} + 918373788279757x^{32} - 2025722841175179x^{31} + 4141678405154077x^{30} - 7863069392075891x^{29} + 13883134774878431x^{28} - 22824971611839946x^{27} + 34978628960492728x^{26} - 50004365561373983x^{25} + 66722398206886406x^{24} - 83127367160315837x^{23} + 96711459871747001x^{22} - 105057149906255752x^{21} + 106520382369921090x^{20} - 100746433329802576x^{19} + 88801393336457818x^{18} - 72855649123022189x^{17} + 55546450812545353x^{16} - 3927390690755884x^{15} + 2568530980157161x^{14} - 15488261636720024x^{13} + 8576797099395441x^{12} - 4340074425776687x^{11} + 1994444842820125x^{10} - 825828427298523x^9 + 305015914916862x^8 - 99169222548261x^7 + 27880041458829x^6 - 6609243354131x^5 + 1272544459004x^4 - 187249389626x^3 + 18791514536x^2 - 968629112x$

$x^{50} - 60x^{49} + 1770x^{48} - 34219x^{47} + 487578x^{46} - 5459915x^{45} + 50034543x^{44} - 385810315x^{43} + 2554404535x^{42} - 14746460937x^{41} + 75125482289x^{40} - 341011791384x^{39} + 1390101790553x^{38} - 5122060131373x^{37} + 17152621486452x^{36} - 5244501368996x^{35} + 14698295809189x^{34} - 378854249719392x^{33} + 900666554177797x^{32} - 1979738477692735x^{31} + 4031929464734980x^{30} - 7621684941656561x^{29} + 13392755217323100x^{28} - 219030581886872x^{27} + 333721978443972x^{26} - 4740615759762502x^{25} + 62821268384716769x^{24} - 77882290736105757x^{23} + 8964600558504704x^{22} - 95183627950206824x^{21} + 96964090608423697x^{20} - 90794033104581729x^{19} + 79182549464255349x^{18} - 64239322567060220x^{17} + 48404836615937201x^{16} - 33808518620233579x^{15} + 2183367353168303x^{14} - 12996812154484384x^{13} + 7103452060911680x^{12} - 3547442534275902x^{11} + 1608856173861531x^{10} - 657490206195397x^9 + 239690301984137x^8 - 76915251291028x^7 + 21335901012750x^6 - 4987043168803x^5 + 945480733568x^4 - 136684969976x^3 + 13431680544x^2 - 674803824x$

$x^{50} - 60x^{49} + 1769x^{48} - 34161x^{47} + 485924x^{46} - 5429003x^{45} + 49608734x^{44} - 381200591x^{43} + 2513564964x^{42} - 14442022312x^{41} + 73176946881x^{40} - 33013856575x^{39} + 1336565824084x^{38} - 4887246487819x^{37} + 1622798564423x^{36} - 49155169616666x^{35} + 136349848401950x^{34} - 3474420298579x^{33} + 81594969605020x^{32} - 1769441758095846x^{31} + 3550933895152601x^{30} - 6605701999466237x^{29} + 11407133902827272x^{28} - 1830679338089421x^{27} + 2732841536365428x^{26} - 37972930406795293x^{25} + 49134125449501746x^{24} - 59216038164915319x^{23} + 66472948221247846x^{22} - 69487280763014941x^{21} + 676112079933081329x^{20} - 61192379001878796x^{19} + 51463766341052977x^{18} - 40168160510889095x^{17} + 2904908631362976x^{16} - 19425441922210309x^{15} + 11981319105886462x^{14} - 6794924380071498x^{13} + 3529738365483165x^{12} - 1671489190099248x^{11} + 717239452033819x^{10} - 27676749133309x^9 + 95099618498032x^8 - 28721240794195x^7 + 7490318664816x^6 - 1645105232061x^5 + 293101653576x^4 - 39859359782x^3 + 3692300568x^2 - 175483464x]$

Figura 6.2.5 – Roteiro para geração de grafo circular com arestas aleatórias e polinômio cromático

```

Passo 1 - Gerar grafos circulares
Cycle= [Complete[0]]
CycleChromPoly= [CompleteChromPoly[0]]
CycleNumberOfColors= [[0]]
for l in range(1,366):
    CycleDict= dict( [ [i, [i+1]] for i in range(l-1) ] )
    CycleDict.update({l-1: [0]})
    CycleDict
    W= Graph(CycleDict, format='dict_of_lists')
    Cycle= Cycle + [W]
    G= W.chromatic_polynomial()
    CycleChromPoly= CycleChromPoly + [G]
    print(G)
    lenG= len(G.coefficients())
    NC= G.coefficients()[-1]
    NumberOfColors= []
    for i in range(0,lenG):
        NumberOfColors= NumberOfColors + [NC]
        H= NC * CompleteChromPoly[lenG-i]
        G= G - H
        NC= G.coefficients()[-1][0]
    NumberOfColors[:::-1]
    print(NumberOfColors)
    CycleNumberOfColors= CycleNumberOfColors + [NumberOfColors]

Passo 2 - Adicionar aresta e Calcular polinômio cromático na forma fatorial
casosn= [50 for i in range(2)]          #[10, 20,30,40,50]
casosp= [0.2 for i in range(5)]        #[0.0, 0.05, 0.1, 0.15, 0.2]
Cycle_Random= []
Cycle_RandomChromPoly= []
Cycle_RandomNumberOfColors= []
for l in casosn:
    for m in range(len(casosp)):
        W= copy(Cycle[l])
        j= 1
        while j <= l*casosp[m]:
            Ini= [ i for i in range(l)]
            shuffle(Ini)
            Fin= [ i for i in range(l)]
            if Ini[0] in Fin: Fin.remove(Ini[0])
            shuffle(Fin)
            if not W.has_edge(Ini[0], Fin[0]):
                W.add_edge((Ini[0], Fin[0]))
                j= j+1
            elif not W.has_edge(Fin[0], Ini[0]):
                W.add_edge((Fin[0], Ini[0]))
                j= j+1
        Cycle_Random= Cycle_Random + [W]
        G= W.chromatic_polynomial()
        Cycle_RandomChromPoly= Cycle_RandomChromPoly + [G]
        lenG= len(G.coefficients())
        NC= G.coefficients()[-1]
        NumberOfColors= []
        for i in range(0,lenG):
            NumberOfColors= NumberOfColors + [NC]
            H= NC * CompleteChromPoly[lenG-i]
            G= G - H
            NC= G.coefficients()[-1][0]
        NumberOfColors[:::-1]
        print(sum(NumberOfColors), NumberOfColors)
        Cycle_RandomNumberOfColors= Cycle_RandomNumberOfColors + [NumberOfColors]

Passo 3 - Plotar grafos gerados
g = []
j = []
# for i in range(len(Cycle_Random): g.append(Cycle_Random[i])
for i in range(len(casosn)):
    n = []
    for m in range(len(casosp)):
        n.append(Cycle_Random[len(casosp)*i + m].plot(layout="circular", vertex_size=20, vertex_labels=False))
    j.append(n)
G = sage.plot.graphics.GraphicsArray(j)
G.show()

Passo 4 - Salvar resultados em um dicionário
Cycle_RandomDict.update({(casosn[0]+casosp[0]):[Cycle_Random, G, Cycle_RandomChromPoly, Cycle_RandomNumberOfColors] })
save(Cycle_RandomDict, "Cycle_RandomDict")
\end{tiny}

```

6.3 Grafos de Redes Sociais

As tecnologias das comunicações se desenvolveram com uma velocidade extraordinária nos dois últimos séculos: o telégrafo, o rádio, o telefone, a televisão até a internet fixa, e hoje móvel. Redes sociais passaram de paroquiais e mercantis para globais. Diferentemente dos casos anteriores, as topologias são variadas e dinâmicas. Um caso de uma rede social muito simples é utilizado como um exemplo de problema dos aniversários. A singeleza facilita a visualização do problema, porém a abordagem adotada pode ser facilmente expandida, sendo as informações sobre a rede e a capacidade de processamento as limitações mais relevantes.

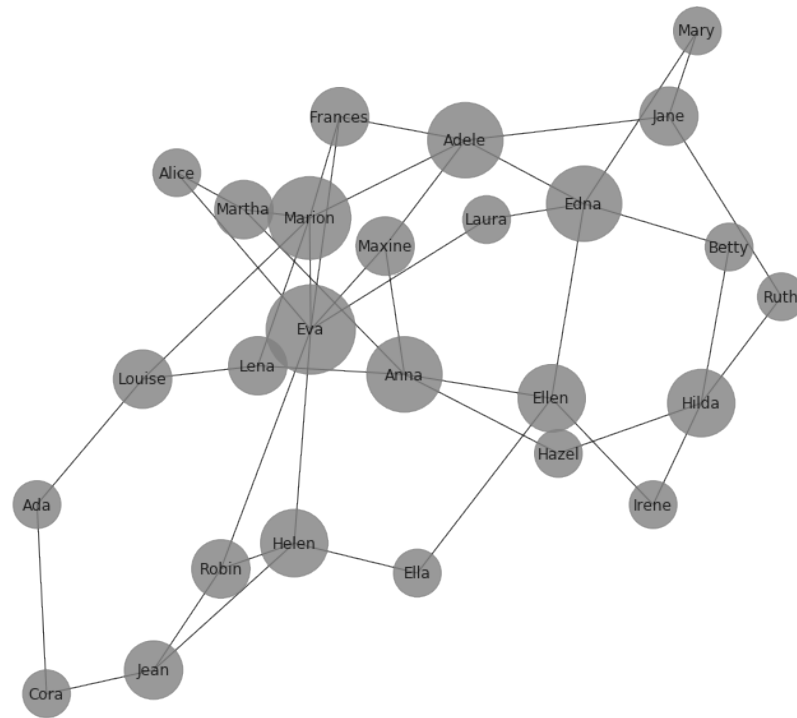
Nooy, Mrvar e Batagelj (2018) apresentam diversos casos de redes sociais, o caso escolhido é fictício e foi desenvolvido por J. L. Moreno para demonstrar como o mapeamento das relações interpessoais permite estudar a estrutura de um grupo, uma área que veio a se consolidar como Sociometria. Um sociograma é uma representação gráfica das relações interpessoais que pode ser representada por grafos, tanto simples como orientados. Basicamente cada pessoa responde a uma ou mais perguntas sobre suas preferências em relação aos outros participantes do grupo. No caso, trata-se de um grupo de moças que compartilham um dormitório de uma escola no Estado de Nova Iorque, EUA. A pergunta foi escolha duas pessoas do grupo que você gostaria de ter ao seu lado na mesa do jantar.

A figura 6.3.1 mostra a representação das respostas na forma de um grafo simples, sendo que as arestas representam as escolhas. Como cada participante escolhe duas colegas, então todas têm pelo menos duas arestas incidentes. Para facilitar a visualização o diâmetro do círculo de cada vértice é proporcional ao seu grau. Uma simples inspeção mostra as pessoas mais populares, como Eva, Marion, Edna e Adele. Também é possível identificar pequenos grupos interligados como Jean, Robin e Helen. O sociograma é um instrumento de apoio, e é utilizado por um profissional que interage diretamente com o grupo.

Este grafo tem 26 vértices e 42 arestas, o que significa uma densidade de $42/325 = 0,129$, se fosse uma árvore seria: $25/325 = 0,077$, ou seja é um grafo esparso.

Um problema dos aniversários para esta rede social é semelhante ao visto na seção 2.3: qual é a probabilidade de que duas ou mais pessoas deste grupo tenham a mesma data de aniversário? As restrições são que estas pessoas sejam parceiras na mesa de jantar.

Figura 6.3.1 – Caso parceiras na mesa do jantar



Fonte: Elaboração própria

Calculando o polinômio cromático resulta:

$$\begin{aligned}
 & +\lambda^{26} - 42 * \lambda^{25} + 856 * \lambda^{24} - 11275 * \lambda^{23} + 107836 * \lambda^{22} - 797625 * \lambda^{21} + 4744775 * \lambda^{20} + \\
 & -23303003 * \lambda^{19} + 96246901 * \lambda^{18} - 338754077 * \lambda^{17} + 1025745399 * \lambda^{16} + \\
 & -2690032484 * \lambda^{15} + 6136814458 * \lambda^{14} - 12207561009 * \lambda^{13} + 21185320036 * \lambda^{12} + \\
 & -32030787860 * \lambda^{11} + 42049453260 * \lambda^{10} - 47659010971 * \lambda^9 + 46240648930 * \lambda^8 + \\
 & -37939106632 * \lambda^7 + 25867909071 * \lambda^6 - 14290622023 * \lambda^5 + 6156168578 * \lambda^4 + \\
 & -1942679508 * \lambda^3 + 399769576 * \lambda^2 - 40263168 * \lambda
 \end{aligned}$$

Calculando a probabilidade de que pelo menos duas colegas tenham a mesma data de aniversário para $\lambda = 365$:

$$Prob(X \geq 1) = 1 - P_{26}(\lambda = 365) = 0,1089$$

Esta probabilidade é muito menor que o problema clássico ($n \simeq 23$ para $P(\lambda) = 1/2$), o que é coerente devido ao grafo ser muito esparso.

Para o polinômio cromático na forma fatorial, resulta que o menor coeficiente diferente de zero é 4797 para $\lambda^{(3)}$, ou seja, é possível dividir o grupo em 3 equipes sem que sejam parceiras na mesa de jantar. São 4797 opções.

7 CONCLUSÃO E RECOMENDAÇÕES

O objetivo principal do trabalho é propor uma generalização do problema dos aniversários quando as restrições de relacionamento são tão complicadas a ponto de ser necessário a construção de uma rede de relacionamentos.

Dos modelos analisados o modelo de von Mises (1939) é o único que pode ser considerado como um modelo genérico e bem estruturado. É lamentável que não haja uma tradução do artigo do alemão para o inglês. A outra contribuição que se destaca na literatura é o modelo dos circuitos quebrados de Whitney (1932). Uma das dificuldades deste assunto é a infinidade de modelos específicos que são de difícil comparação para pesquisadores que não são especialistas na área.

Da revisão da bibliografia clássica fica evidente a dificuldade de se modelar este tipo de problema a partir das ferramentas da análise combinatória, o modelo de caixas e bolas. A aplicação do método da inclusão exclusão nestes casos é extraordinariamente trabalhoso. Os polinômios cromáticos fazem esta operação de forma estruturada e confiável, porém podem exigir tempos de processamento também extraordinariamente longos.

Outro aspecto importante desta discussão é que a álgebra dos polinômios é bastante conhecida e são existentes sistemas simbólicos que permitem automatizar uma parte relevante do trabalho. A mudança de bases possibilita que se escolha a forma mais intuitiva ou simples de tratar um problema e possibilita sua análise de diferentes pontos de vista. O caso das árvores é exemplo claro, onde um argumento simples de construção do polinômio cromático pode transformá-lo nas bases de potência ou fatorial.

Apesar do limitado conhecimento da Matemática Discreta pelo autor, fica evidente as inúmeras possibilidades de analogias, ou bijeções na linguagem matemática, entre suas diversas áreas, o que demanda uma formação mais abrangente dos profissionais da área técnica para desenvolver soluções de problemas importantes, muitos deles urgentes.

Nas áreas pesquisadas pelo autor, não foi identificada abordagem que trate a generalização proposta para o problema dos aniversários em redes de relacionamento de forma mais eficaz que a aplicação de métodos de contagem por polinômios cromáticos.

Um objetivo secundário é a análise das alternativas para a o cálculo dos polinômios cromáticos. Esta mesma dificuldade é comum em toda a literatura que associa características de grafos e polinômios, sendo os cromáticos uma parte importante, porém muitos outros são estudados e utilizados na solução de problemas.

Dos sistemas de cálculo disponíveis atualmente, como as linguagens de programação e ambientes de desenvolvimento, apenas um grupo limitado delas oferece um ambiente que possibilita a experimentação, que é essencial na formação e no desenvolvimento dos pesquisadores e profissionais ligados a área de simulação de sistema complexos.

A opção pelo Sagemath em Linux foi abordada ao longo do trabalho. Analisando retrospectivamente, foi uma escolha essencial para o desenvolvimento do trabalho e do aprendizado do autor, através da experimentação em pequenos problemas que juntos dão forma ao trabalho realizado. A facilidade das bibliotecas permite que se experimente diferentes alternativas com um pequeno esforço.

Os algoritmos conhecidos e disponíveis são relativamente simples, o que significa que há oportunidades interessantes na área. Por exemplo, métodos híbridos de remoção/contração e de adição/contração podem ser indicados para grafos esparsos com áreas de concentração, como "hubs". A solução de polinômios cromáticos demanda operações simples e que facilmente permitem sua separação em operações distintas, o que é típico de qualquer método que se baseie no princípio da inclusão exclusão. Técnicas de processamento paralelo podem contribuir para elevar os limites atuais a solução deste tipo de problema.

Este trabalho propõe uma importante generalização do problema dos aniversários, porém é evidente que este assunto é muito mais abrangente do que foi possível para o autor desenvolver.

Outra possibilidade para tratar de grafos de tamanhos maiores, é o uso de métodos que implementam simulações de Monte Carlo de cadeias de Markov, quais devem ser comparados com os analíticos para identificar os limites de efetividade dos algoritmos.

O problema dos aniversários embute uma dificuldade adicional quando se trata de coincidências múltiplas, que se revelam, por exemplo, nas florestas de circuitos quebrados como formadoras dos coeficientes de polinômios na base de potência.

Para o limitado conhecimento do autor, este assunto está aberto, tanto para os métodos clássicos como para os probabilísticos.

Vale lembrar que o trabalho se concentrou em parte do aspecto probabilístico das coincidências, os estudos dos vieses cognitivos e das decisões parcialmente "racionalis" também estão em desenvolvimento.

A própria definição de coincidência é um problema aberto na comunidade científica e tão comum na nossa vida cotidiana.

REFERÊNCIAS

- ABRAMSON, M.; MOSER, W. O. J. More birthday surprises. *The American Mathematical Monthly*, v. 77, n. 8, p. 856–858, 1970.
- ALDOUS, D. Another conversation with persi diaconis. *Statistics Science*, v. 28, n. 2, p. 269–281, 2013.
- BAXTER, R. J. *Ecatly solved models in statistical mechanics*. Suffolk, UK: Springer, 1982.
- BERRESFORD, G. C. The diameter of a cicle plus a random matching. *Mathematics Magazine*, v. 53, n. 5, p. 286–288, Nov 1980.
- BHATTACHARYA, B. B.; DIACONIS, P.; MUKHERJEE, S. Universal limit theorems in graph coloring problems with connections to extremal combinatorics. *The Annals of Applied Probability*, v. 27, n. 1, p. 337–394, Feb 2017.
- BOLLOBAS, B.; CHUNG, F. R. K. The diameter of a cicle plus a random matching. *SIAM J. Disc. Math.*, v. 1, n. 3, p. 328–333, Aug 1988.
- BONDY, J.; MURTY, U. *Graph Theory*. New York: Springer, 2008.
- CAMBRAY, J. *Sincronicidade: natureza e psique num universo interconectado*. Petrópolis: Vozes, 2013.
- CHAO, C.-Y.; WHITEHEAD, E. G. On chromatic equivalence of graphs. In: ALAVI, Y.; LICK, D. R. (Ed.). *Theory and Applications of Graphs*. [S.l.]: Springer-Verlag, 1978. p. 121–131.
- CHIA, G. L. A bibliography on chromatic polynomials. *Discrete Mathematics*, v. 172, n. 1–3, p. 175–191, 1997.
- DASGUPTA, A. The matching, birthday and the strong birthday problem: a contemporary review. *Journal of Statistical Planning and Inference*, v. 130, n. 1–2, p. 377–389, 2005.
- DIACONIS, P. Theories of data analysis: from magical thinking through classical statistics. In: HOAGLIN, D. C.; MOSTELLER, F.; TUKEY, J. W. (Ed.). *Exploring Data Tables, Trends, and Shapes*. Hoboken, NJ: John Wiley, 1985. p. 1–36.
- DIACONIS, P. Five stories for Richard. In: HERSH, P.; LAM, T.; PYLYAVSKYY, P.; REINER, V. (Ed.). *The Mathematical Legacy of Richard P. Stanley*. Providence: American Mathematical Society, 2016. p. 131–140.
- DIACONIS, P.; HOLMES, S. A bayesian peek into feller volume i. *Sankhya: The Indian Journal of Statistics, Series A*, v. 64, p. 820–841, 2002.
- DIACONIS, P.; MOSTELLER, F. Methods for studying coincidences. *Journal of American Statistical Association*, v. 84, n. 408, p. 853–861, Dec 1989.
- DONG, F. M.; KOH, K. M.; TEO, K. L. *Chromatic polynomial and Chromaticity of Graphs*. 3a. ed. Singapore: World Scientific Publishing, 2005.
- ESTRADA, E.; KNIGHT, P. A. *A First Course in Network Theory*. Oxford: Oxford University Press, 2015.
- FADNAVIS, S. A generalization of the birthday problem and the chromatic polynomial. p. 1–17, Sept 2015. ArXiv: 1105.0698v3. Disponível em: [⟨arxiv.org/abs/1105.0698⟩](https://arxiv.org/abs/1105.0698).
- FALK, R. Judgment of coincidences: Mine versus yours. *The American Journal of Psychology*, v. 102, n. 4, p. 477–493, 1989.

- FALK, R. A closer look at the notorious birthday coincidences. *Teaching Statistics*, v. 36, n. 2, p. 41–46, Jun 2014.
- FALK, R. On coincidences in academia and fiction. *Archives of Psychology*, v. 1, n. 2, p. 1–10, Nov 2017.
- FELLER, W. *An Introduction to Probability Theory and Its Applications, Vol. I*. 3a. ed. New York: John Wiley, 1968.
- FEOFILOFF, P.; KOHAYAKAWA, Y.; WAKABAYASHI, Y. *Uma Introdução Sucinta à Teoria dos Grafos*. 2011. Disponível em: www.ime.usp.br/~pf/teoriadosgrafos/texto/TeoriaDosGrafos.pdf.
- GRIFFITHS, T. L.; DANIELS, D.; AUSTERWEIL, J. L.; TENENBAUM, J. B. Subjective randomness as statistical inference. *Cognition*, v. 103, p. 85–109, 2018.
- GRIFFITHS, T. L.; TENENBAUM, J. B. Randomness and coincidences: Reconciling intuition and probability theory. In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. [s.n.], 2001. v. 23, p. 1–6. Disponível em: escholarship.org/uc/item/2cw487hh.
- GRIFFITHS, T. L.; TENENBAUM, J. B. From mere coincidences to meaningful discoveries. *Cognition*, v. 103, n. 2, p. 180–226, 2007.
- HAGGARD, G.; PEARCE, D. J.; ROYLE, G. Computing tutte polynomials. *ACM Transactions on Mathematical Software*, ACM, v. 37, n. 3, p. 24:1–24:17, 2010.
- HAND, D. J. *The Improbability Principle – Why Coincidences, Miracles, and Rare Events Happen Every Day*. New York: Scientific American, 2014.
- HOCKING, R. L.; SCHWERTMAN, N. C. An extension of the birthday problem to exactly k matches. *The College Mathematics Journal*, MAA, v. 17, n. 4, p. 315–321, 1986.
- JUNG, C. G. Synchronicity: An acausal connecting principle. In: _____. *The Interpretation of Nature and the Psyche*. New York: Pantheon Books, 1955. Original: *Naturerklärung und Psyche*. Zürich: Rascher Verlag, 1952.
- KAHNEMAN, D. *Rápido e Devagar: Duas Formas de Pensar*. Rio de Janeiro: Objetiva, 2011.
- LEWIS, R. M. R. *A Guide to Graph Colouring: Algorithms and Applications*. Cham: Springer International, 2015.
- LIMA, A. M. d. *Algoritmos exatos para o problema da coloração de grafos*. Dissertação (Mestrado) — Universidade Federal do Paraná, 2017.
- LOERINC, B. M. *Computing Chromatic Polynomials for Special Families of Graphs*. New York: Courant Institute of Mathematical Sciences, 1980.
- MAZUR, J. *Acaso: Como a Matemática Explica as Coincidências da Vida*. Rio de Janeiro: Casa da Palavra, 2016.
- MCKINNEY, E. H. Generalized birthday problem. *The American Mathematical Monthly*, MAA, v. 73, n. 4, p. 385–387, 1966.
- MONAGAN, M. A new edge selection heuristic for computing the tutte polynomial of an undirected graph. p. 1–11, Sept 2012. ArXiv: 1209.5160v1. Disponível em: arxiv.org/abs/1209.5160.
- MOSTELLER, F. Understanding the birthday problem. *The Mathematics Teacher*, v. 55, n. 5, p. 322–325, 1962.
- NIJENHUIS, A.; WILF, H. S. *Combinatorial Algorithms for computers and calculators*. 2nd ed. ed. New York: Academic Press, 1978.

- NOOY, W. de; MRVAR, A.; BATAGELJ, V. *Exploratory Social Network Analysis with Pajek*. 3rd ed. ed. Cambridge: Cambridge University Press, 2018.
- NUNNIKHOVEN, T. S. A birthday problem solution for nonuniform birth frequencies. *The American Statistician*, v. 46, n. 4, p. 270–274, 1992.
- PAH, C. H. An application of catalan numbers on cayley tree of order 2: Single polygon counting. *Bulletin of the Malaysian Mathematical Sciences Society*, v. 31, n. 2, p. 175–183, 2008.
- PAULI, W. The influence of archetypal ideas on the scientific theories of Kepler. In: _____. *The Interpretation of Nature and the Psyche*. New York: Pantheon Books, 1955. Original: *Naturerklärung und Psyche*. Zürich: Rascher Verlag, 1952.
- PETERSON, I. *The Jungles of Randomness: a mathematical safari*. [S.l.]: John Wiley, 1998.
- READ, R. C. An introduction to chromatic polynomials. *Journal of Combinatory Theory*, v. 4, p. 52–71, 1968.
- READ, R. C. *An improved method for computing the chromatic polynomials of sparse graphs*. [S.l.]: University of Waterloo, Faculty of Mathematics, 1987.
- RUST, P. F. The effect of leap years and seasonal trends on the birthday problem. *The American Statistician*, v. 30, n. 4, p. 197–198, 1976.
- STANLEY, R. P. *Catalan Numbers*. New York: Cambridge University Press, 2015.
- TVERSKY, A.; KAHNEMAN, D. Judgment under uncertainty: Heuristics and biases. *Science*, American Association for the Advancement of Science, v. 185, n. 4157, p. 1124–1131, 1974.
- VON MISES, R. Ueber Aufteilungen - und Bezetzungswahrscheinlichkeiten. *Revue de la Faculté des Sciences de l'Université d'Istanbul N. S.*, v. 4, p. 145–163, 1939.
- WANG, F. X.; CHEN, G. Complex networks: Small-world, scale free and beyond. *IEEE circuits and systems magazine*, IEEE, v. 3, n. 1, p. 6–20, 2003.
- WEISSTEIN, E. W. *Chromatic Polynomial*. 2019. Disponível em: mathworld.wolfram.com/ChromaticPolynomial.html.
- WHITNEY, H. A logical expansion in mathematics. *Bulletin of the American Mathematical Society*, v. 38, n. 8, p. 572–579, 1932.

ANEXO A – ROTINA DE VON MISES (1939)

A seguinte função implementa o cálculo de número de possibilidades de ocupação de n bolas distintas em Lambda caixas distintas, conforme von Mises (1939):

```
def CompleteDistr(DistrCaixas, Lambda):
#
# von Mises 1939 Ocupação de n bolas distintas em Lambda caixas distintas
# Dados:
#   DistrCaixas = lista com número de caixas com s bolas, s= 1,2, ...
#   Lambda = número de caixas disponíveis
# Restrições:
#   Número de caixas = Lambda, logo Número de Caixas vazias =
#                                     Lambda - sum(DistrCaixas) >= 0
#   Número de bolas =
#       sum( (i+1)*DistrCaixas[i] for i in range(DistrCaixas) ) > 1
#
# Calcula a distribuição das bolas por tipo de caixa (s bolas)
#
DistrBolas= []
for i in range(len(DistrCaixas)):
    DistrBolas= DistrBolas + [(i+1)*DistrCaixas[i]]
#
# Insere número de caixas vazias na posição 0 da lista DistrCaixas
#
NumCaixasOcup= sum(DistrCaixas)
DistrCaixas= [Lambda-NumCaixasOcup] + DistrCaixas
#
# T1= multinomial(DistrCaixas)
#
aux1= 1
for i in range(len(DistrCaixas)):
    aux1*= factorial(DistrCaixas[i])
T1= factorial(Lambda)/aux1
#
# T2= permutação (fatorial) do total de bolas
#
NumBolas= sum(DistrBolas)
T2= factorial(NumBolas)
#
# T3= produto dos arranjos das bolas nos grupos
#
aux4= 1
for i in range(1,len(DistrCaixas)):
    aux4*= factorial(i)^DistrCaixas[i]
T3= aux4
# Número de possibilidades da ocupação dada das bolas
#                                     nas Lambda caixas disponíveis
TT= T1*T2/T3
return TT
```

ANEXO B – ROTINA DE HOCKING E SCHWERTMAN(1986)

O seguinte programa implementa o cálculo de número de possibilidades de ocupação de n bolas distintas em Λ caixas distintas, sendo que k caixas têm duas bolas e o restante uma bola, conforme 4.2. Utiliza a função proposta por von Mises (1939) no anexo A :

```
#
# Hocking e Schwertman 1986:
#     Ocupação de n bolas distintas em Lambda caixas distintas
#     com exatas Nduplas de 1 a 8 (coincidências simples)
# Dados:
#     DistrCaixas = lista com número de caixas com s bolas, s= 1, 2
#     Lambda = número de caixas disponíveis
# Restrições:
#     Número de caixas = Lambda, logo Número de Caixas vazias =
#                                     Lambda - sum(DistrCaixas) >= 0
#     Número de bolas =
#         sum((i+1)*DistrCaixas[i] for i in range(DistrCaixas) ) = x > 1
#     Número de caixas com duas bolas = de 1 a 8
#

import matplotlib.pyplot as plt

var('x, Lambda') # permite cálculo simbólico com variáveis x e
                 # Lambda (número de bolas e de caixas)

Lambda= 365
#
#     Ciclo do número de duplas
#
Nduplas= 8
T=[]
Ngraf= [[0]]
Tgraf= [[0]]
for i in range(Nduplas+1):
#
#     Constrói a distribuição de caixas, calcula o número de bolas e
#     função T(x,Lambda) das possibilidades
#     i caixas com duas bolas e (x-2*i) caixas com uma bola: total de x bolas
#
    DistrCaixas= [x-2*i,i]
    NumBolas= sum((j+1)*DistrCaixas[j] for j in range(len(DistrCaixas)))
    T+= [CompleteDistr(DistrCaixas, Lambda)]
#
#     Constroi a lista para gráfico com limite Limgraf menor que Lambda= 365
#
    Limgraf= 120
    Lxgraf= []
    Lgraf= []
```

```
for j in range(2*Nduplas,Limgraf):
    Lxgraf+= [j]
    Lgraf+= [((T[i]).subs(x=j,Lambda=365)/Lambda^j)]
Ngraf+= [Lxgraf]
Tgraf+= [Lgraf]
#
# Plota as curvas de Prob x Número de Bolas para todas as Nduplas
#
plt.xlabel("Numero de Bolas")
plt.ylabel("Probabilidade de k duplas")
plt.title("Probabilidade de k duplas (Lambda= 365)")
for i in range(1,Nduplas+1):
    plt.plot(Ngraf[i],Tgraf[i],label = '%s duplas'%i)
plt.legend()
plt.show()
```

ANEXO C – ROTINA DE MCKINNEY (1966)

O seguinte programa implementa o cálculo de número de possibilidades de ocupação de n bolas distintas em λ caixas distintas, adicionadas uma a uma até que uma caixa atinja k bolas, conforme 4.3. Utiliza a função proposta por von Mises (1939) no anexo A :

```
#
# McKinney 1966: Problema dos aniversários generalizado:
#           qual o menor número de bolas necessárias,
#           adicionadas uma a uma,
#           para alcançar k coincidências com probabilidade 1/2
# Restrições:
#   Número de caixas = Lambda,
#           logo Número de Caixas vazias = Lambda - sum(DistrCaixas) >= 0
#   Número de bolas =
#           sum( (i+1)*DistrCaixas[i] for i in range(DistrCaixas) ) > 1
#   Número de caixas com k bolas = 1
#

import time

var('x, Lambda') # n = x
Lambda= 365
McKinneyCasos=[[2,21],[3,86]] # ,[4,185]]
PMcKinney= []
#
# r= 2, 3, 4 . Retire o comentário para processar r=4
#           (tempo de longo processamento - alguns dias).
#
for Caso in range(len(McKinneyCasos)):
    r= McKinneyCasos[Caso][0]
    n= McKinneyCasos[Caso][1]
    Prob= 0

    while Prob < 5/10 and n < 366:
        start = time.time()
        n+= 1 # número de bolas

#
# Gera todas situações que atendem as restrições:
#   n_i < r, i= 1, 2, ..., r-1 e somatorio i * n_i= número de bolas
# Cada restrição é uma lista N[i],
#           onde cada elemento é o número de caixas com i bolas
#
        naux= IntegerListsLex(n, min_part=1, max_part=r-1 ,
                               min_slope=-(r-1), max_slope=0)

#
# Transforma a lista gerada em IntegerListsLex para o formato [n_1,n_2,...,n_r-1]
#
        N= []
```

```

TN= 0
Taux= len(naux)*[0]
for i in range(0,len(naux)):
    N.append((r-1)*[0])
    for j in range(0,len(naux[i])):
        N[i][naux[i][j]-1]+= 1
    aux= 0
    for j in range(1,len(N[i])):
        aux+= (j+1) * N[i][j]
    print(N[i])
    T(Lambda)= CompleteDistr(N[i], Lambda)
    Taux[i]+= T
for i in range(0,len(naux)):
    TN+= Taux[i]
Pr(Lambda)= TN / Lambda ^ n
Prob= 1 - Pr.subs({Lambda:365})
end = time.time()
Duration= end - start
PMcKinney+= [r, n, Prob, Duration, Pr(Lambda), TN]
#
# Imprime resultados
#
    print(r, n, Prob.n(digits=4), round(Duration,4), TN)
    print
print
print

```

ANEXO D – ROTINA DE ÁRVORES DE CAYLEY

O seguinte programa implementa a geração e o cálculos de grafos completos e estrela, que são árvores. O polinômio cromático de uma árvore conexa só depende do seu número de vértices.

Os polinômios de grafos completos são utilizados como divisores dos outros polinômios para realizar a mudança da base de potência para base fatorial.

```

import time
#
#  Geração de Grafos Completos
#
#  Polinômios cromáticos são utilizados como divisores de polinômios
#  na mudança da base de potência para fatorial
Complete= []
for i in range (0,366):
    W= graphs.CompleteGraph(i)
    Complete= Complete + [W]
Complete
#
#  Polinômio cromático
#
var(x)
function('WCP', nargs=1)
CompleteChromPoly= []
CompleteDuration= []
for i in range (0,366):
    start = time.time()
    WCP= sage.arith.misc.falling_factorial(x, i).full_simplify()
    CompleteChromPoly= CompleteChromPoly + [WCP]
    end = time.time()
    Duration= end - start
    CompleteDuration= CompleteDuration + [Duration]
    print(i, round(CompleteDuration[i-1], ndigits=8))
#
#  Geração de grafos estrela (árvores)
#
#  Mudança de base de potência para base fatorial
#
import time
Star= [Complete[0]]
StarChromPoly= [CompleteChromPoly[0]]
StarNumberOfColors= [[0]]
StarDuration= [0]
for l in range(1,366):
    start = time.time()
    W= graphs.StarGraph(l-1)
    Star= Star + [W]
    var(x)
    G= x * (x-1)^(l-1)

```

```
G= G.full_simplify()
StarChromPoly= StarChromPoly + [G]
print(G)
lenG= len(G.coefficients())
NC= G.coefficients()[-1][0]
NumberOfColors= []
for i in range(0,lenG):
    NumberOfColors= NumberOfColors + [NC]
    H= NC * CompleteChromPoly[lenG-i]
    G= G - H
    NC= G.coefficients()[-1][0]
NumberOfColors[::-1]
print(NumberOfColors)
StarNumberOfColors= StarNumberOfColors + [NumberOfColors]
end = time.time()
Duration= end - start
StarDuration= StarDuration + [Duration]
print(1, round(StarDuration[1-1], ndigits=8))
```


ANEXO E – ROTINA DE GRAFOS DE BOLLOBAS-CHUNG

O seguinte programa implementa a importação de dados da rede social, o desenho do grafo da rede e a solução de um problema dos aniversários com as restrições de parceria entre as integrantes do grupo.

```
#
# Importa biblioteca de funções aleatórias
#
import random
from random import shuffle
#
# Gera um grafo circular e adiciona arestas aleatórias
#
# Verifica se a aresta é existente
#
l=100
GraphCR= copy(Cycle[l])
seed()
j= 1
while j <= 4:
    Ini= [ i for i in range(l)]
    shuffle(Ini)
    Fin= [ i for i in range(l)]
    if Ini[0] in Fin: Fin.remove(Ini[0])
    shuffle(Fin)
    if not GraphCR.has_edge(Ini[0], Fin[0]):
        GraphCR.add_edge((Ini[0], Fin[0]))
        j= j+1
    elif not GraphCR.has_edge(Fin[0], Ini[0]):
        GraphCR.add_edge((Fin[0], Ini[0]))
        j= j+1
GraphCR.plot(layout= "circular")

#
# Histograma de graus
#
GraphCR.degree_histogram()

Out: [0, 0, 93, 6, 1]
#
# Polinômio cromático
#
GraphCR.chromatic_polynomial()
#
# Gera batelada de casos
#
casosp= [0.10 for i in range(10)]; casosp
#
```

```
# Plota os grafos gerados
#
g = []
j = []
# for i in range(len(Cycle_Random): g.append(Cycle_Random[i])
for i in range(len(casosn)):
    n = []
    for m in range(len(casosp)):
        n.append(Cycle_Random[len(casosp)*i + m]
                .plot(layout="circular", vertex_size=20, vertex_labels=False))
    j.append(n)
G = sage.plot.graphics.GraphicsArray(j)
G.show() # long time
```

ANEXO F – ROTINA DE REDES SOCIAIS

O seguinte programa implementa a importação de dados da rede social, o desenho do grafo da rede e a solução de um problema dos aniversários com as restrições de parceria entre as integrantes do grupo.

```
#
# Rede Social: Caso parceiras da mesa de jantar
#
# Importa bibliotecas e rede Pajek Dormitory.net
#
import networkx as nx
import matplotlib.pyplot as plt

G_Dorm= nx.Graph(nx.read_pajek("Dormitory.net"))
G_Dorm= G_Dorm.to_undirected()
#
# Cria grafo no formato Sagemath para cálculo do polinômio cromático
#
Dorm= Graph(G_Dorm)
#
# Imprime número de vértices, número de arestas e densidade
#
print len(Dorm.vertices()), len(Dorm.edges())
print 2*len(Dorm.edges())/(len(Dorm.vertices()*(len(Dorm.vertices())-1))
#
# Imprime lista de graus e arestas
#
print
print Dorm.degree()
print
print Dorm.vertices()
print
print G_Dorm.edges()
#
# Imprime grafo layout spring
#
plt.figure(figsize=(10,9))
pos= nx.spring_layout(G_Dorm)
nx.draw_networkx(G_Dorm, pos, alpha=0.8,
node_size=[700*G_Dorm.degree(v) for v in G_Dorm],node_color= 'gray')
plt.axis('off')
plt.tight_layout()
#
# Imprime grafo layout circular
#
plt.figure(figsize=(10,9))
pos= nx.circular_layout(G_Dorm)
nx.draw_networkx(G_Dorm, pos, alpha=0.8,
```

```

node_size=[700*G_Dorm.degree(v) for v in G_Dorm],node_color= 'gray')
plt.axis('off')
plt.tight_layout()
#
# Calcula a polinômio cromático e a probabilidade Lambda= 365
#
var (x)
DormChromPoly= Dorm.chromatic_polynomial()
DormProb= DormChromPoly/x^len(Dorm)
Prob= 1 - (DormProb(x=365)).n()
print
print DormChromPoly
print
print 'Prob=', Prob.n(digits=4)
#
# Calcula polinômio cromático base fatorial
#
var(x)
G= DormChromPoly
lenG= len(G.coefficients())
NC= G.coefficients()[-1]
NumberOfColors= []
for i in range(0,lenG):
    NumberOfColors= NumberOfColors + [NC]
    H= NC * falling_factorial(x,lenG-i).full_simplify()
    G= G - H
    NC= G.coefficients()[-1][0]
NumberOfColors[::-1]
print
print(NumberOfColors)
print
i= lenG-1
while NumberOfColors[i]== 0:
    i-= 1
print lenG-i, NumberOfColors[i]
print

Resultados Sagemath

26 42
42/325

[3, 4, 2, 3, 2, 2, 5, 3, 3, 2, 2, 2, 3, 4, 5, 7, 2, 2, 3, 3, 3, 5, 4, 2, 6, 2]

[u'Ada', u'Adele', u'Alice', u'Anna', u'Betty', u'Cora', u'Edna', u'Ella', u'Ellen',
~u'Eva', u'Frances', u'Hazel', u'Helen', u'Hilda', u'Irene', u'Jane', u'Jean',
~u'Laura', u'Lena', u'Louise', u'Marion', u'Martha', u'Mary', u'Maxine', u'Robin',
~ u'Ruth']

[(u'Louise', u'Lena'), (u'Louise', u'Marion'), (u'Louise', u'Ada'),
~(u'Ellen', u'Edna'), (u'Ellen', u'Ella'), (u'Ellen', u'Irene'),
~(u'Ellen', u'Anna'), (u'Ada', u'Cora'), (u'Jane', u'Ruth'),

```

~(u'Jane', u'Adele'), (u'Jane', u'Mary'), (u'Martha', u'Marion'),
 ~(u'Martha', u'Alice'), (u'Martha', u'Anna'), (u'Mary', u'Edna'),
 ~(u'Anna', u'Maxine'), (u'Anna', u'Lena'), (u'Anna', u'Hazel'),
 ~(u'Lena', u'Marion'), (u'Ruth', u'Hilda'), (u'Cora', u'Jean'),
 ~(u'Ella', u'Helen'), (u'Hazel', u'Hilda'), (u'Frances', u'Eva'),
 ~(u'Frances', u'Marion'), (u'Frances', u'Adele'), (u'Helen', u'Eva'),
 ~(u'Helen', u'Jean'), (u'Helen', u'Robin'), (u'Marion', u'Eva'),
 ~(u'Marion', u'Adele'), (u'Eva', u'Maxine'), (u'Eva', u'Alice'),
 ~(u'Eva', u'Laura'), (u'Eva', u'Robin'), (u'Betty', u'Edna'),
 ~(u'Betty', u'Hilda'), (u'Jean', u'Robin'), (u'Maxine', u'Adele'),
 ~(u'Edna', u'Laura'), (u'Edna', u'Adele'), (u'Hilda', u'Irene')]

$$\begin{aligned}
 &x^{26} - 42x^{25} + 856x^{24} - 11275x^{23} + 107836x^{22} - 797625x^{21} \\
 &+ 4744775x^{20} - 23303003x^{19} + 96246901x^{18} - 338754077x^{17} \\
 &+ 1025745399x^{16} - 2690032484x^{15} + 6136814458x^{14} - 12207561009x^{13} \\
 &+ 21185320036x^{12} - 32030787860x^{11} + 42049453260x^{10} - 47659010971x^9 \\
 &+ 46240648930x^8 - 37939106632x^7 + 25867909071x^6 - 14290622023x^5 \\
 &+ 6156168578x^4 - 1942679508x^3 + 399769576x^2 - 40263168x
 \end{aligned}$$

Prob= 0.1089

[1, 283, 35706, 2660681, 130785878, 4487051972, 110906729734, 2010387334417,
 ~26959933692837, 268006879220402, 1967816768252329, 10576961364883877,
 ~41021313228735057, 112446673973737132, 211767293376291817, 263688558560229255,
 ~206060632995027638, 94009591292878860, 22576835441861114, 2443261552320236,
 ~92849248023553, 798146354598, 635519379, 4797, 0, 0]

3 4797