

精神科相談サイトのテキストマイニング

山形東高等学校 2年 小野晶子 指導教員 佐藤勝治先生

序論 なぜこの研究をしたのか？ →医療診断支援への応用の可能性

いつも使っているTwitterで心の病の悩みを投稿する人がいる→どんな悩みを抱えているのか知りたい【先行研究】

人工知能を精神科医療に応用する取り組みの中で、患者の書いた文章を分析し、特定の単語の使用頻度からその患者の重症度を測定するシステムが開発中である。(1,2)

先行研究では患者から独自に取得したデータを使用しているため、データを効率的に集めるのが難しい→自分がネット上のデータを使えば大量に効率よく集められるのではないかな？

もし、単語の傾向に性差、年齢差があるなら患者の属性に対応した治療の一助に？



研究対象と方法

【用語解説：うつ病、双極性障害、統合失調症】

うつ病：抑うつ気分が続く

双極性障害：抑うつ状態と躁状態（気分高揚）を繰り返す。躁うつ病という呼称が有名

統合失調症：幻覚、妄想が見れる

①ウェブサイト「Dr林のこころと脳の相談室」(http://kokoro.squares.net/)内の文章を使用「うつ病」「双極性障害」「統合失調症」カテゴリーに分類されているもの(2013.5.3-2019.5.5投稿)を読む

②回答者（回答者が判断を避けている場合は相談者の主治医）が「うつ病（双極性障害、統合失調症）の可能性が高い」と認定している投稿を抽出

③本人からの相談か周囲からか、年齢、性別で分類→結果①-1

④文章内の名詞、動詞、形容詞を抽出し、出現頻度解析を行った。株式会社ユーザーローカル提供のテキストマイニングツールを使用

④-1「時間の経過や薬の量、その単位を表す単語」「薬を表す単語」「薬の具体的な名称を含む相談」の数をそれぞれ集計→結果①-2

④-2共起ネットワークを行った→結果②
薬の処方や具体的な量についての単語が多く集まっている部分を○で囲んだ
「思う」など感情を表す表現が多く集まっている部分を○で囲んだ

結果①-1

年	性	男性	女性	計
10代	0	3	3	
20代	9	13	22	
30代	6	7	13	
40代	11	7	18	
50代	2	1	3	
計	28	31	60(不詳1)	

図-うつ病患者本人からの相談件数(相談者の属性別)

年	性	男性	女性	計
30代	0	4	4	
40代	0	2	2	
50代	3	1	4	
60代以上	2	8	10	
不詳	2	3	5	
計	7	18	26(不詳1)	

図-うつ病患者の周囲からの相談件数(患者の属性別)

年	性	男性	女性	計
10代	0	1	1	
20代	3	9	12	
30代	4	14	18	
40代	6	8	14	
50代	0	2	2	
60代~	0	1	1	
計	13	35	48	

図-双極性障害患者本人からの相談件数(相談者の属性別)

年	性	男性	女性	計
10代	0	1	1	
20代	0	0	0	
30代	0	1	1	
40代	4	1	5	
50代	0	1	1	
60代~	0	1	1	
計	4	7(不詳2)	11	

図-双極性障害患者の周囲からの相談件数(患者の属性別)

年	性	男性	女性	計
10代	3	25	28	
20代	34	70	104	
30代	16	23	39	
40代	10	10	21(その他1)	
50代	2	1	3	
計	65	129	195	

図-統合失調症患者本人からの相談件数(相談者の属性別)

年	性	男性	女性	不詳	計
10代	9	8	1	18	
20代	20	19	0	39	
30代	22	28	0	50	
40代	9	19	1	29	
50代	5	14	0	19	
60代~	2	25	0	27	
不詳	22	22	5	49	
計	89	135	7	231	

図-統合失調症患者の周囲からの相談件数(相談者の属性別)

結果①-2

	うつ病	双極性障害	統合失調症
相談件数(件) ...a	86	57	426
単語の総数(語) ...b	17240	16805	79355
時間の経過や薬の量、その単位を表す単語の数(語) ...c	953	1170	1940
c/b (%)	5.47	6.96	2.44
薬を表す単語の数(語) ...d	614	625	1262
d/b (%)	3.52	3.72	1.59
薬の具体的な名称を含む相談(件) ...e	57	38	100
e/a (%)	66.28	64.41	23.47
d/e (語/件)	10.27	16.45	12.62

【用語解説：共起ネットワーク】

共起ネットワーク：文章中に同時に出現しやすい単語同士の関係を図にすること

【共起ネットワーク図の見方】

出現数が多い単語ほど円が大きくなる

共起の程度が強い(同時に出現しやすい)ほど単語間の線が太くなる

赤:動詞 青:名詞 緑:形容詞 灰:感動詞

【用語解説：テキストマイニング】

文章(テキスト)を処理してパターンを発見しようと言う手法。分析には専用ツールを用いて、特定の単語の出現率を計算したり、同じ分にはどうい単語とどうい単語が一緒に存在するかという共起性を導き出したりすることや大量の定性データを比較的短時間に処理することが可能となっている。(広告用語辞典)

考察①

全体として薬の話が多い

全体の195/571件中 34.15%

うつ病と双極性障害に限れば95件/145件中 65.52%(過半数)

a、「精神科＝話を聴く」というイメージ(いのちの電話など?)、しかしそれだけでなく薬による治療が重要

b、患者の関心「薬は効いているか?効きすぎ(副作用が出ている)ではないか?」

薬に関する内容を含む相談1件あたりの単語数(語/件)

双極性障害16.45 > 統合失調症12.62 > うつ病10.77

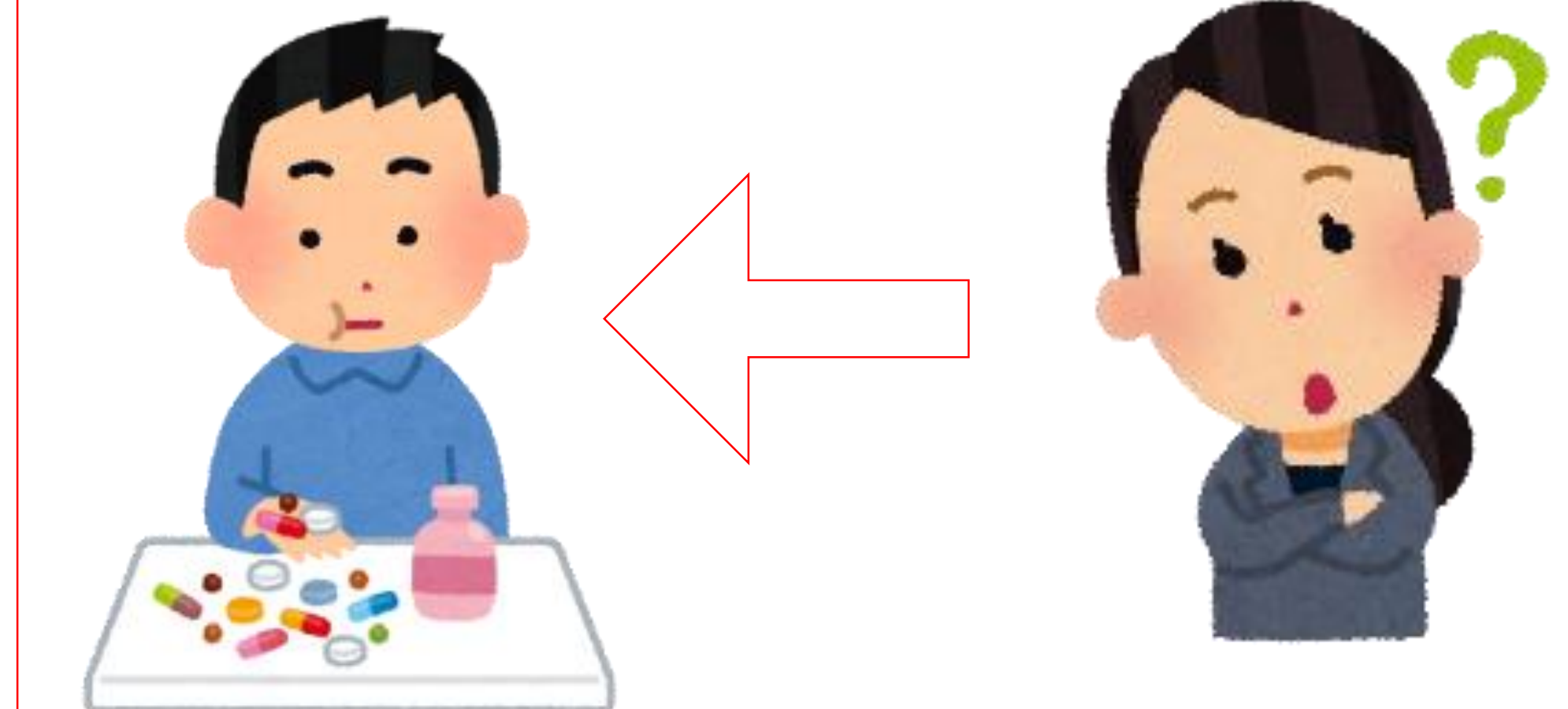
→既に薬をもらっている患者の場合、**双極性障害、統合失調症患者は相談する際に薬のことを多く話す傾向がある。**

なぜ?

うつ病はきれいに治り、薬がいなくなることがある

統合失調症は薬がいなくなることにはあるが、判断が慎重になされる

双極性障害は薬を飲み続けていく必要がある(3,4) →双極性障害、統合失調症は薬への関心が高い



考察②

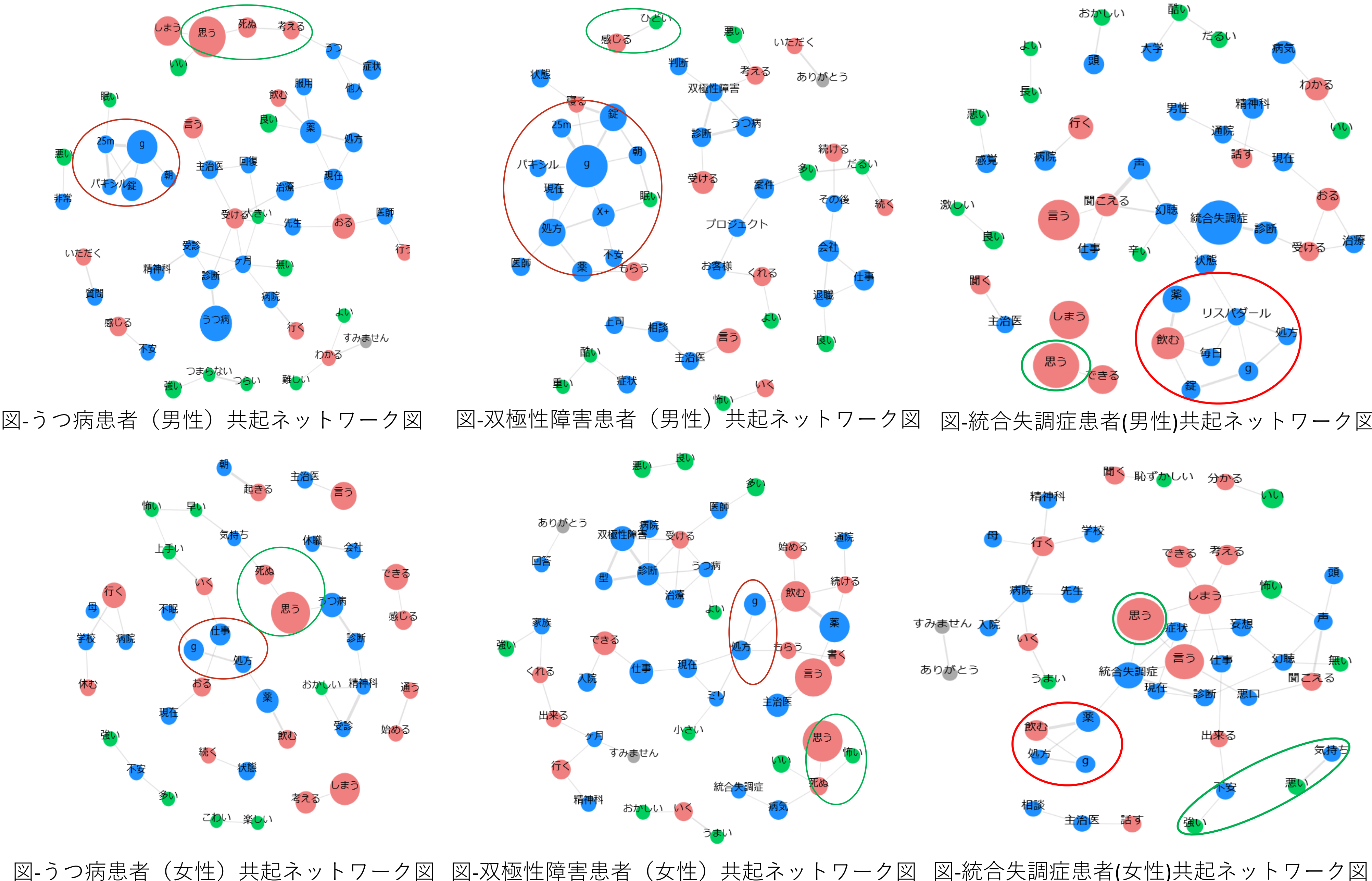
具体的な処方に関する内容は、男性に多く女性に少ない
思いや感情に関する内容は、女性に多く男性に少ない
なぜ???わかりませんでした...



今後の展望

- ・他の病気(など)ではどのような傾向があるか調べたい
- ・今回は2013.5.3以降のデータを調査したが、サイト自体は1997.4.16から運営→古いデータも調べたい
- ・リサーチクエスト「患者はどのようなことが気になっているのか」→相談理由のデータを取りたい
- ・相談に男女差が生まれる理由を考察したい
- ・医療相談に役立てられればよいなあ

結果②



参考文献

- 1.情報通信技術や機械学習を活用した精神疾患重症度評価への取り組み 岸本 泰士郎,リョウ コクケイ,工藤 弘毅,吉村 道孝,田澤 雄基,吉田 和生 2017,情報管理,vol60,no8
- 2.言葉から精神症状の特徴を抽出 心の病の予防に新しい手掛かり(岸本 泰士郎氏インタビュー) 2017,Webサイト「AI時代と科学研究の今」(https://www.jst.go.jp/kisoken/jyonetsu/interview/h29/kishimoto.html)
- 3.双極性障害(躁うつ病)-疾患の詳細-専門的な情報-メンタルヘルス-厚生労働省 (https://www.mhlw.go.jp/kokoro/specialty/detail_bipolar.html)
- 4.統合失調症-疾患の詳細-専門的な情報-メンタルヘルス-厚生労働省 (https://www.mhlw.go.jp/kokoro/specialty/detail_into.html)

謝辞

ウェブサイト「Dr林のこころと脳の相談室」(http://kokoro.squares.net/)の運営をされている林公一先生にサイト内の文章の使用許可をいただきました。株式会社ユーザーローカル提供のテキストマイニングツール(https://textmining.userlocal.jp/)を分析に使用させていただきました。佐藤勝治先生にこの研究をご指導いただきました。この場を借りてお礼を申し上げます。