

## コミックからの登場人物相関図の作成

京極 亮太<sup>†</sup> 上田 洋<sup>‡</sup> 村上 晴美<sup>§</sup>

<sup>† §</sup>大阪市立大学大学院創造都市研究科

<sup>‡</sup>大阪市立大学大学院工学研究科

### 1. はじめに

近年、国内外で日本コミックの人気の高まっている。過去に出版されたコミックから、現在も連載されているコミック、海外のコミックを含めると膨大な数となり、どのようなコミックがあるのか個人が把握するのは難しい。そこで本研究では、コミックの特徴を表現して個人の好みのコミックを探す手助けとして、登場人物の相関図に着目する。相関図は人手で作成されることが多く、作り手によって表現が異なる。また、一つのコミックで多いものでは100巻を超えるため、人手で作成するのは労力を要する。本研究では簡単にコミックの登場人物相関図を作成する方法を提案する。既存の相関図作成の研究として[1][2]などがあるが、これらはテキストから言語処理により作成するものである。本研究ではコミックからテキストを用いずに相関図を作成する。

本稿の構成は以下のとおりである。2節で提案手法について述べる。3節では提案した手法の実用例について述べる。4節では評価実験について述べる。

### 2. 提案手法

本研究では、「出現頻度が高い登場人物(以下人物)と、共に出現する人物の情報を利用することが相関図作成に有効である」という仮説をたて、コミックの特徴であるコマ割に着目した手法を提案する。

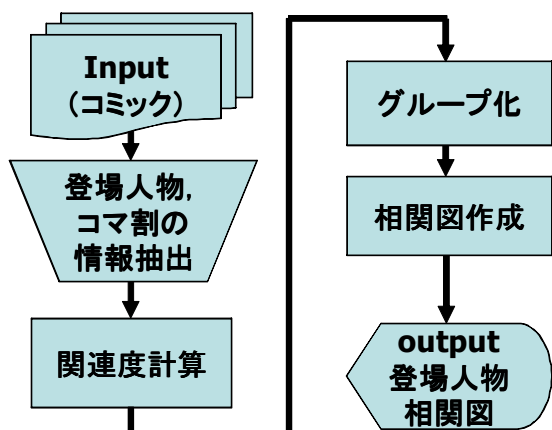


図1: 提案手法の概要

提案手法の概要は以下のとおりである。まず、コマ内に出現する名前のある人物を人手で抽出する。次に、人物間の関連度を計算する。関連度はコマ内の出現情報と、次のコマへの出現情報を用いる。人物の出現頻度と人物間の関連度を用いて、人物のグループ化を行い、このグループを使用して相関図を作成する。図1に提案手法の概要を示す。

#### 2.1 登場人物・コマ割情報の抽出

コミックから、コマと、コマ内に出現する名前のある登場人物を人手で抽出する。ページや話数などの情報も抽出する。抽出支援用のツールを作成して使用した。

#### 2.2 関連度計算

コミックは、コマ内に景色や人物を描くことでストーリーを展開する。このコマを使用して関連度計算を行う。

##### 2.2.1 コマ内関連度

一つのコマに複数の人物が登場する場合、このコマ内の人物間に関連度を与える。人物が少ないほど関連度を高くする。関連度  $R_1$  はコマ内の登場人数  $n$  のとき、

$$R_1 = 1/(n-1) \quad (1)$$

とする。人物が5人以上の場合、 $n$ は5とする。

##### 2.2.2 コマ外関連度

コミックはコマによってストーリーを進めるため、あるコマに登場する人物と、その次のコマに登場する人物間に関連度を与える。コマ内関連度と同様に、次のコマに登場する人物が少ないほど関連度を高くする。あるコマに複数の人物がいる場合、その中から1人ずつ、次のコマへの関連度を計算する。次のコマの内、対象の人物以外の人数を  $m$  として、関連度  $R_2$  は以下のとおりとする。

$$R_2 = 0.5/m \quad (2)$$

人物が5人以上の場合、 $m$ は5とする。

コマ外関連度計算は、1話を区切りとして処理を行う。これは、話数を跨ぐと関連度の低い人物に関連度を与えてしまう可能性があるためである。

##### 2.2.3 両関連度の結合

コマ内関連度  $R_1$  とコマ外関連度  $R_2$  を結合して本手法の関連度とする。結合した関連度  $R$  の式は、

$$R = R_1 + R_2 \quad (3)$$

とする。

### 2.3 グループ化

相関図作成に使用するグループを作成する（図 2 参照）。

- Step1** 各人物を出現頻度順に並べ、人物毎に関連する人物を関連度順に並べる（表 1 参照）。
- Step2** 各人物毎の関連度の差を調べ、最も差の大きい箇所にて区切り、左側の人物を「該当人物の上位グループ」とする。
- Step3** 各「該当人物と該当人物の上位グループ」に含まれる人物が相互に含まれている場合と同じグループとする（表 2 参照）。

図 2：グループ化

表 2：グループ化の例

グループ	該当人物	関連人物
1	セル	ベジータ
2	クリリン	トランクス
3	クリリン	18号
4	悟空	悟飯
5	悟空	天津飯
6	18号	16号
7	天津飯	ピッコロ

図 4 は、提案手法で作成したドラゴンボール第 32 巻の登場人物相関図である。ドラゴンボール全体の主人公は「悟空」であるが、第 32 巻の中心人物は「セル」であり、「セル」を中心とした相関図が作成された。

### 2.4 相関図作成

人物の出現頻度、人物間の関連度、グループを使用して相関図を作成する（図3参照）。

- Step1** 最も出現頻度の高い人物 A を中心に配置する。
- Step2** 人物 A のグループ中の人物を下部に配置する。
- Step3** 人物 A のグループ以外で出現頻度の高い上位人物 5 名を B として配置する。ただし、このとき、
  - Step3.1** 人物 B にグループが存在する場合、グループ中の人物を B の周囲に配置する。
- Step4** Step1～Step3 で配置されていない人物を、出現頻度の高い順に配置する。
- Step5** 関連度が 10 以上ある人物間を矢印で結ぶ。

図 3：相関図作成

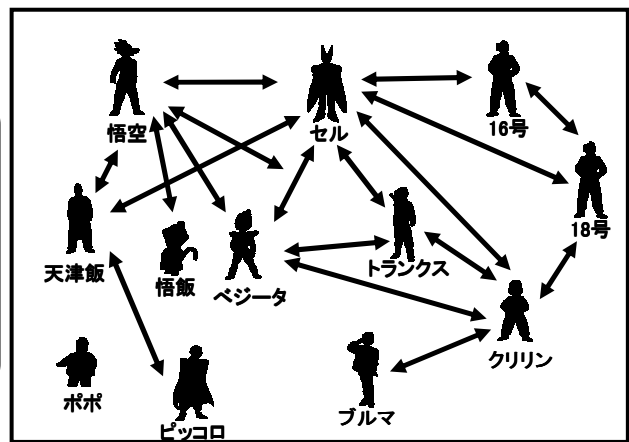


図 4：相関図作成の実行例

## 3. 実行例

表 1 は、ドラゴンボール第 32 巻のデータセットから、出現頻度の高い上位 6 名の人物と関連度の高い上位 6 名の人物を示したものである。たとえば、最も出現頻度の高い「セル」の「ベジータ」との関連度は 172.5 である。表 2 はグループ化の例である。「セル」と「ベジータ」はグループであり、「クリリン」と「トランクス」、「クリリン」と「18号」はグループである。

相関図では、Step 1 で最も出現頻度が高い「セル」、Step 2 で「セル」のグループ 1 内の「ベジータ」が配置される。Step 3 でグループ 2-7 内の人物が配置される。Step 4 でまだ配置されていない人物が配置される。

表 1：関連度の例（出現頻度の上位 6 人物）

該当人物	頻度	1	2	3	4	5	6
1 セル	266	ベジータ	トランクス	16号	クリリン	18号	悟空
		172.5	43.1	24.0	23.7	15.7	14.3
2 ベジータ	196	セル	トランクス	悟空	クリリン	16号	18号
		172.5	28.2	10.4	10.2	8.3	6.3
3 クリリン	115	トランクス	セル	18号	ブルマ	16号	ベジータ
		52.6	43.1	20.1	14.0	11.4	10.2
4 トランクス	115	クリリン	セル	ベジータ	悟空	ブルマ	16号
		52.6	43.1	28.2	13.2	8.1	5.5
5 悟空	69	悟飯	天津飯	セル	トランクス	ベジータ	ピッコロ
		43.8	15.0	14.3	13.2	10.4	8.9
6 18号	53	16号	クリリン	セル	ベジータ	トランクス	天津飯
		35.8	20.1	15.7	6.3	3.1	0.5

(注) 下段の数字は関連度

## 4. 評価実験

ドラゴンボール第 32 巻（全 126 ページ、12 話）を使用した。被験者 5 名に第 32 巻を読ませた後、作成した相関図（図 4）を提示して 5 段階で評価（5：非常にあっている、4：ややあっている、3：どちらともいえない、2：やや違う、1：違う）させた。実験結果は平均 4.4 であり、本手法の有効性を示唆している。事後インタビューの結果、「セル」を中心配置することへの妥当性を確認した。

## 5. おわりに

コミックのコマ割に着目し、人物の共出現情報を利用して登場人物の相関図を作成する手法を提案した。被験者 5 名に対する相関図の 5 段階評価は 4.4 で、本手法の有効性を示唆している。現在評価実験を継続中である。

### 参考文献

- [1] 馬場こづえ, 藤井敦: 小説テキストを対象とした人物情報の抽出と体系化, 言語処理学会第 13 回年次大会発表論文集, 574-577, 2007.
- [2] 後藤 淳, 八木伸行, 相澤彰子, 関根 聡: 照応解析を利用した放送番組からの登場人物の相関図生成, 2008 年度人工知能学会全国大会 (第 22 回) 論文集, 2008.