

創作者と人工知能: 共作実現に向けた創作過程とメタデータ付与 4 コマ漫画ストーリーデータセット構築

Creators and Artificial Intelligence:
Enabling Collaboration with Creative Processes and Meta-data for Four-scene Comic Story Dataset

上野 未貴^{*1}

Miki Ueno

^{*1}豊橋技術科学大学, 情報メディア基盤センター

Toyohashi University of Technology, Information and Media Center

Recently, researches in the field of Artificial Intelligence(AI) have become widely developed so that computational methods to analyze and build creative contents have attracted a notable following. Especially, computational methods as deep learning, which is one of the methods of machine learning for content creation has rapidly developed; i.e. caption generation, automatic colorization, and sketch drawings. On the other hand, creativity in humans is an intellectual activity meaning that it is really hard to simulate the whole creative process as computational methods. In this field, the number of data for machine learning is too small because creative processes often contain precious and sensitive data. In this research, we propose and construct a four-scene comics story dataset with creative process and meta-data toward collaboration between "Creators and Artificial Intelligence".

1. はじめに

人工知能技術の発展により, 創作物を計算機が理解したり, 自動生成する [1] ことに強い関心が集まっている. 画像キャプション生成 [2] や自動彩色, スケッチ生成 [3] など, 創作に関わる特に深層学習を利用した技術に近年大きな発展が見られる. 一方で, 人の創作は高次の知的作業であり, 計算機上に創作過程全体を具現化することは極めて難しい. この分野において, 創作物が貴重で繊細な情報をしばしば含むため, 機械学習向けのデータが非常に少ない. 本研究では, 「創作者と人工知能」の共作に向け, 創作過程とメタデータを有する 4 コマ漫画ストーリーデータセットを構築する.

2. 基本方針:研究姿勢と本研究の目的

2.1 研究姿勢

五感から入力を得て, 感じ考えたことを絵や言葉で表して創作物として出力する全体機構 [4] は, 現在, 人にしかできない. 人工知能分野の技術が発展しても, 感情や経験, 背景知識なしには, フレーム問題 [5] に代表されるように, 機構の大部分の実現は計算機では困難と考えられてきた. 問題を定式化するため, ストーリーを持つ創作物はある種, 現実世界との関連を保って状況構築した有限世界の好題材と考えられる. 広範には, 工学は最終的には人の役に立つ技術手法を研究する学問とされる. ストーリーを有し, 言語, 画像, 音などの多くの要素を含む創作物であるアニメの創作現場環境には課題が多く, 人が一番時間をかけたい作業に時間をかけられないという問題が発生することがある. 例えば「旅立つ」ストーリーを描く際, 各種漫画を参考に, 主人公の表情, 台詞などについて, 新たな表現を試みるため, 類似資料が存在していないか調べるコストがかかる. 小説, 漫画, 絵本, アニメなどの, 創作場面ではアナログな資料管理が主流であり, デジタルデータを残す文化が少ない. また人同士の対面のコミュニケーションで進行する制作が多く, 従来, 計算機科学との親和性は高くない. 一方, デ

ジタル技術の発展により, 積極的に技術活用する現場もある. 工学系の研究者として, 技術発展のために, 創作物を貴重なデータとしてもみなす必要がある. 一方で, 創作物は長い時間をかけて創作者が創り上げた作品であり現場の意見を丁寧に拾い扱う必要がある. 以上の考えから, 著者はこれまでから, 創作に関わる人へ意見を聞き, 要素技術 [6] と応用手法^{*1}の研究を進めてきた. 人が自由に利用したい情報技術を選択して利用する, という姿勢で創作支援を目指す.

2.2 本研究の目的

1. 漫画のストーリーを機械学習で扱うためのパターン定義

創作の中でも着目する過程に異なりがある. 著者はストーリー理解に特に着目している. ストーリーを計算機で扱う研究は小説や漫画などの形態を問わず困難とされる.

言語と画像のマルチモーダルな媒体である漫画は, 芸術的, 心理学的見地だけでなく, 工学研究 [7] においても注目される. 特に, 著者は, 漫画の中でも, 形状と長さの統一, ストーリーが明瞭という観点から, 日本で流通が多い 4 コマ漫画を対象に取る. n コマ漫画を計算機に生成 [8] する研究を進め, 深層学習を 4 コマ漫画に適用し感情識別・順序識別する要素技術を提案した. 一方, 現状許諾を得られる既存作品のストーリー分析には制約が存在する. 本研究で, 計算機と親和性が高い 4 コマ漫画パターンを提案し, 新たなデータセットの構築指針を示す.

2. 創作過程を有した創作物のデータ構築と応用方法

解決すべきタスクの細分化が困難, 実現場での応用方法のイメージ不足, 権利関係や既存作品から創作過程と作者の付与したメタデータを得ることが難しい. そのため, 計算機上で機械学習に利用できるストーリーを有するデータが少なく, 分野発展への壁がある. 本研究では, 創作者と計算機の共作を目指し, 作者の創作過程と読者の読解過程のストーリーに関わるメタデータが付与された 4 コマ漫画ストーリーデータセットを構築する. また, 想定タスクと現場応用方法を述べる.

連絡先: 上野 未貴, 豊橋技術科学大学, 情報メディア基盤センター, 〒441-8580 愛知県豊橋市 天伯町雲雀ヶ丘 1-1,
E-mail: ueno@imc.tut.ac.jp

^{*1} 創作者と人工知能分野研究者のコラボレーション:
<https://www.tut.ac.jp/docs/PR180222.pdf>

3. 分野動向:創作物の工学系研究

本分野における、最終的な研究目的は、技術的に共通する部分もあるが、多くの場合、自動生成と創作支援、解析と生成に大別される。比較的ルールが厳密に定義でき、近年、深層学習の興隆から飛躍的な進歩を遂げるゲームプレイプログラム AlphaGo などとは状況が異なる。創作過程の要件を細分化して定義することは極めて困難であり、現状、全自動で明らかに人を凌駕する技術があるとは言えない。

3.1 創作物の種別と特徴

創作物に関する工学系の研究と特徴を述べる。

3.1.1 漫画

オブジェクト認識、自動コマ割レイアウト、レビュー分析、吹き出し推定、線画補正、ネタバレ解析、読解教材応用などの研究 [7] が知られる。

3.1.2 小説

ストーリーの扱いが難しく、全文自動生成は困難だと考えられている。しかし、深層学習による部分的な文章生成が進展し、創作支援の立場を取る研究も増加している。

3.1.3 アニメ

アニメは創作過程において、自然言語、静止画、動画、音楽と扱うデータが多く、各作品の創作に関わる人数が多い [1]。創作過程の絵コンテは、形状が同一のコマ画像と台詞や状況説明の自然言語から成り、本研究で対象とするコマ漫画と類似する。カメラワークや尺などの情報がある。

3.1.4 絵本

対話的な絵本の自動生成 [9] が知られる。連続する絵でストーリーを表すだけでなく、各ページを鑑賞する時間が長く、一枚の絵の中に複数のストーリーが存在することが多い。ストーリーを解析する際にはストーリーの領域分割が必要と考えられる。見開きの2ページで、絵のページと、文字だけのページを独立させる場合と、見開きの2ページともに、絵の中に白抜きで文字を入れる場合が存在する。

3.2 漫画の従来データセット

漫画の研究のデータセットとして、国内ではプロの漫画 109 冊の全ページの画像を電子化したデータセット MANGA109[10] が知られる。ジャンルや年代が異なる漫画が広く含まれており、異なるタッチの漫画オブジェクト認識の研究等に向く。4コマ漫画は5冊含まれるが、作者や話のテーマが大きく異なる。漫画の形態や流通状況は国別に異なる部分もあるが、海外でも漫画は工学系の題材として着目されている。フランスでは、eBDtheque [11] というデータセットが知られる。漫画 100 ページに対し、台詞やコマ位置などのメタデータを付与したもののだが、ジャンルに偏りが見られる。いずれも、ストーリーに関わるデータを作成することは主目的とされていない。

3.3 従来データセットの問題点:ストーリーの観点

ストーリーは、俳句、古典落語などに典型的に見られるように、型に嵌めること、型を破ることで生まれる。4コマ漫画は決められたコマ内にストーリーを画像と言語で表す型があり、計算機科学の良質な題材である。一方、同一オブジェクトでもタッチが異なること、オリジナリティの観点から同一プロットを異なる作者が書くことは稀である。そのため、データを集められたとしても、共通部分を見つける難しさがあり、ストーリーの解析のための深層学習の適用に課題がある。一方、ストーリーの定量化が難しいことから、素性設計が困難である。

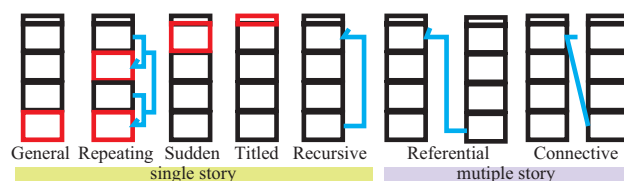


図 1: 4 コマ漫画ストーリーパターン 7 種
注:特にオチのあるコマを赤色で、読む順と関連を水色で表す。

4. 4 コマ漫画構造ストーリーパターンの定義

問題解決のため著者は 4 コマ漫画のストーリーパターンを提案している。コマ構造について、アメリカで有名な 4 コマ漫画ピーナッツに対し認知心理学的な読解過程の研究 [12] が存在するが、本研究とは機械学習適用を主目的としない点で異なる。図 1 に、本研究で再整理した 4 コマ漫画のストーリーパターン 7 種を示し、以下、図の左から順に対応させ説明する。

1. 一般 標準的な起承転結。原則的に 4 コマ目にオチがある。
2. 繰返し 1-2 コマ間の類似事象が 3-4 コマにも起きる *2。
3. 出オチ 1 コマ目におかしな絵が描かれてオチがある *3
4. タイトルオチ 最後にタイトルを見返してオチがわかる。
5. 再帰 4 コマ目から 1 コマ目に戻り話として成立する。
6. 参照 1 つ以上前の話の続きの話。
7. 連続 連続した 4 コマを 2 話 並べて 8 コマで話となる。

新聞や、広告の一部に掲載される 4 コマ漫画もあるが、流通する漫画雑誌や単行本では 1 ページに 2 話の 4 コマ漫画が並ぶことが通例である。日本では特に、連載の間隔が短い。ストーリー 4 コマ漫画と呼ばれる各話で話が短く完結しているが、複数話で全体の大きな話を持つ、連載などに向いた形態が多い。ストーリー 4 コマ漫画では、単一話と複数話という 2 種を定義し、構造を考える必要がある。

単一話とは、各話で完結する話である。複数話とは、2 話以上で関係を持つ話である。

ストーリーパターンを、大きく構造と内容に分けて定義する。網羅性に課題があり限定的ではあるが、広義には、構造はタイトルと各コマのオチの有無の組合せ 2⁵ 種のパターンで表せると考えた。中でも、使用頻度が高いか、視覚的に特徴的な組合せには、制作現場で呼称が存在する。

頻度が高いパターンは、「一般」であり、他に視覚的に特徴的なものは「出オチ」である。

5. 4 コマ漫画ストーリーデータセット構築

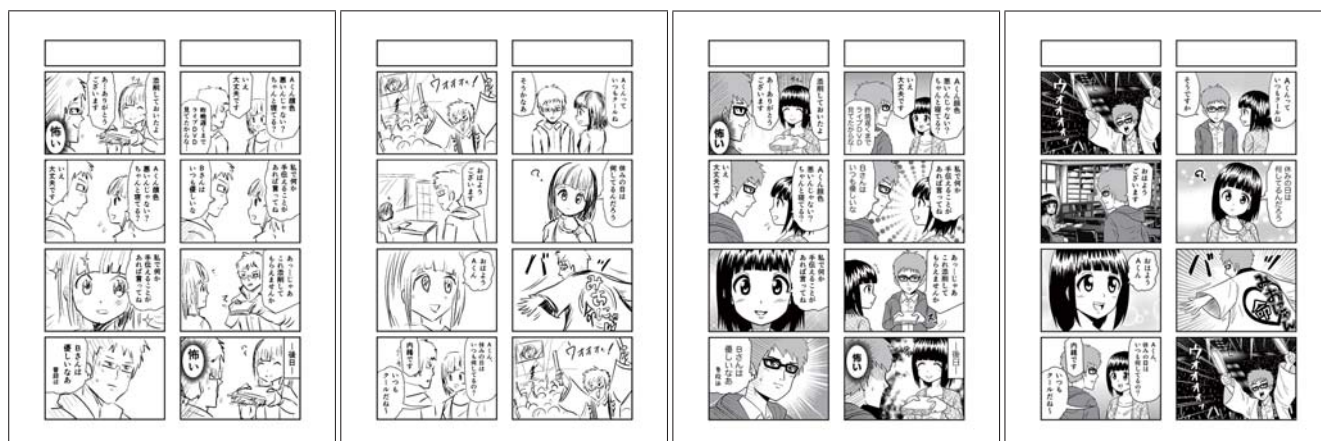
本研究では、上記 2 種の構造に係るストーリーパターン（以下、構造ストーリーパターン）に着目し、内容に関わるストーリーパターンをメタプロットとして、なるべく固有名詞を含まない抽象化したテキストで表した。

5.1 構築方針

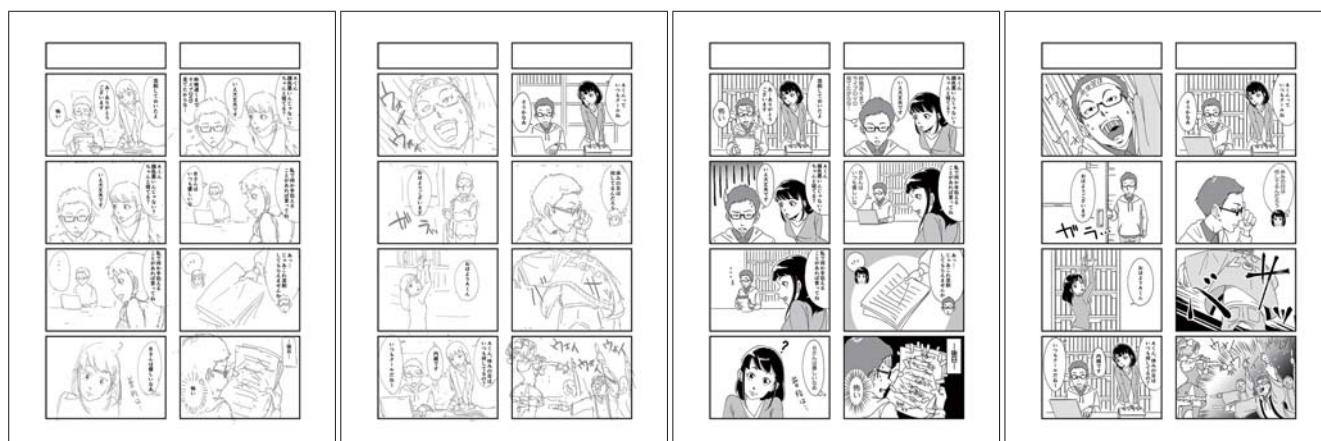
キャラクタ情報、舞台情報、10 個のメタプロットとプロットを作成し、2 つの構造ストーリーパターンを指定して、20 個のシナリオを作成し、5 人の異なるタッチ:ギャグ、少年、青年、少女、萌え、のプロ作者が計 100 作品の漫画を制作した。

*2 3-4 コマ目の方が強調された事象が起きる。

*3 4 コマ目にもオチがあることが多い。



©作画: 浦田カズヒロ (シナリオ: 株式会社スポマ 播村 早紀/豊橋技術科学大学 上野 未貴)



©作画: 鈴木市規 (シナリオ: 株式会社スポマ 播村 早紀/豊橋技術科学大学 上野 未貴)

図2: 第1, 2話, 2構造, 2作者のラフと清書

注:日本では原則として, 漫画は各話に関して各コマ内で右から左, コマ間で上から下, ページ内で右から左へ読む。

■第一話-構造 1

[1] B「Aくんっていつもクールね」A「そうかな」

[2] B「心の声(休みの日は何してるんだろ)」

[3] A「法被を着る」

[4] アイドルのコンサートで法被を着てサイリウムを振る A

■第一話-構造 2

[1] アイドルのコンサートで法被を着てサイリウムを振る A

[2] 場所: 研究室挨拶をして研究室に入ってくる A「おはようございます」

[3] B「おはよう Aくん」

[4] B「Aくん、休みの日はいつも何してるの?」A「内緒です」B「いつもクールだね」

■第二話-構造 1

[1] B「Aくん顔色悪いんじゃない?ちゃんと寝てる?」A「いえ大丈夫です」A心の声(昨晚遅くまでライブDVD見てたかな...)

[2] B「私で何か手伝えることがあれば言ってね」A(Bさんはいつも優しいな)

[3] A「あっ...じゃあこれ添削してもらえませんか」Bに原稿を渡す A

[4] ナレ: 一後日一赤と付箋だらけの原稿が返却される A心の声(怖い)

■第二話-構造 2

[1] Bからびっしり赤と付箋だらけの原稿を渡される A

B「添削しておいたよ」A「あ...ありがとうございます」A心の声(怖い)

[2] B「Aくん顔色悪いんじゃない?ちゃんと寝てる?」A「いえ大丈夫です」

[3] B「私で何か手伝えることがあれば言ってね」

[4] A心の声(Bさんは優しいなあ)書き文字: 普段は

図3: 第1, 2話のシナリオ

注:作者へ提示したものになるべく沿い, 本稿用に調整した。

5.2 創作過程

本データセットの創作過程は以下のとおりである。合わせて作業の担当を示す。各過程から得られる各情報は自然言語情報もしくは画像情報として保持した。

1. キャラクタおよび舞台設定を決定 [著者]

- 共通の主人公キャラクタ2人の情報共有
- 共通の舞台設定の情報共有

2. メタプロットを作成 [著者]

3. メタプロットからプロットを作成 [著者]

4. プロットからシナリオを作成 [制作会社と著者]

5. シナリオからラフを作成 [各漫画作者]

6. ラフから清書 [各漫画作者]

5.3 レイヤ情報およびメタデータの付与

本データセットを作成するにあたり可能な限り創作に係る情報を保持してデータを付与した。

レイヤ情報: コマ枠/各キャラクタの表情に係る顔部位/各キャラクタの顔部位以外/各吹き出し/各台詞/各背景

メタデータ付与: 計算機上で扱いやすいフォーマットを考慮し, 各作者と読者の双方が以下のメタデータを付与した。

- キャラ単位の感情7種: 基本6感情とニュートラル
- オチを説明するキャプション: 文字数を指定

6. データ具体例と考察

全体の舞台とキャラ設定, 第1, 2話のデータを具体例として示す。図2に2作者のラフ画像と清書を示す。表1にメタプロットとプロット, 図3にシナリオを示す。

表 1: 第 1, 2 話のメタプロットとプロット

話数	メタプロット	プロット
1		A キャラ紹介: 見た目は <クール> だが実は <熱中しているもの <アイドル> がある>
2	見た目は <1> だが実は <2>	B キャラ紹介: 見た目は <可愛い> が実は <怖い一面がある>

注: シナリオ制作用のデータ。漫画作者へは提示しない。本稿用に内容変更せず表記調整した。

舞台: 大学の研究室。主人公はキャラクタ A, B の学部生と大学院生の二人。人工知能分野の感性情報系の研究をしており異なるテーマを持つ。

キャラ特徴: 外見として、髪、目、身長、体系、服装、内面として、年齢、学年、兄弟構成、性格、等を設定。

共通シナリオには多様な表現方法がある。例えば 2 作者間において、第 1 話、構造 1, 1 コマ目、1, 4 コマ目の構図が類似する。また、第 1 話、構造 1, 1 コマ目では、構図は類似するがカメラワークが異なる。第 1 話、構造 1, 2 コマ目ではキャラクタの視点が異なる。

異なる作者の共通クラスの単一イラスト画像の学習手法 [3] は発展している。一方で、特にプロが一般的に同じシナリオに関する漫画を書くことは稀である。本研究では、創作過程を保持し、共通シナリオに対する漫画データを構築することで、多様な表現、例えば、構図やカメラアングルなど、ストーリーの演出に関わる表現について、また、吹き出しや背景効果などの漫画特有表現の考察が可能である。

担当編集や演出が担うような計算機手法の構築が考えられる。創作時に経験豊富なプロに相談するのはもちろんだが、相談が難しい場合に、気軽に助言を求められるパートナーとしての計算機手法の創出を目指す。

7. 展望: 4 コマ漫画ストーリーデータセットの実験タスク例と要素技術の応用例

本データセットを用いた、深層学習による 4 コマ漫画データセットの実験タスクの例を示す。以下、著者提案の 2 つの 4 コマ漫画ストーリー理解の要素技術を例に取る。

感情識別: Convolutional Neural Network (CNN) を用いた 4 コマ漫画のコマの感情識別 [4]

順序識別: CNN を用いた 4 コマ漫画のコマの順序識別 [6]

本データセットを用いて構築した要素技術の応用例として、プロットの創作支援や、創作中の関連資料検索などが考えられる。4 コマ漫画以外の漫画やクリップアートなどの画像に対して提案された、オブジェクト識別やキャラクタ識別にも用いることを想定して構築した。

8. まとめ

本研究では、漫画の機械学習の研究において、研究初期段階に十分な共通データを揃えることが難しい、という従来の既存漫画を用いることによる問題に対し、あらかじめ、作者と読者の両面を考慮したメタデータを、創作過程の言語情報と画像情報、および、レイヤ情報を残し、権利関係に配慮して、4 コマ漫画データを構築する仕組みを提案した。4 コマ漫画ストーリーデータセットは「創作者と人工知能」の研究発展のため、順次公開を予定する。今後の課題は以下の通りである。

- データの増加
- 実応用タスクに関する意見聴取
- 深層学習フレームワークの構築

謝辞

本研究は、JST, ACT-I の支援を受けたものである; 課題名「深層学習による 4 コマ漫画のストーリー解析用データセット及びフレームワークの開発」。

本稿で述べたデータセットの制作に対し、株式会社スポマ 代表取締役 白井 英明 氏、同社 播村 早紀 氏、天野 聡彦 氏、本稿への掲載漫画を作成して下さい、浦田カズヒロ 氏、鈴木市規 氏へ謝意を示す。

参考文献

- [1] Miki Ueno, Kiyohito Fukuda and Naoki Mori, "Can Computers Create Comics and Animations?", *Computational and Cognitive Approaches to Narratology*, pp. 164–190, 2016.
- [2] Kuniaki Saito, Andrew Shin, Yoshitaka Ushiku, and Tatsuya Harada, "DUALNET: DOMAIN-INVARIANT NETWORK FOR VISUAL QUESTION ANSWERING", *The 18th IEEE International Conference on Multimedia and Expo*, 2017
- [3] David Ha and Douglas Eck, "A Neural Representation of Sketch Drawings". *ArXiv e-prints*, 2017.
- [4] 上野 未貴, 井佐原 均. "漫画内の特徴的要素が与えるストーリーの印象についての検討". 人工知能学会, No. 215-OS-08b-4in2, 2016.
- [5] John McCarthy and Patrick J. Hayes, "Some philosophical problems from the standpoint of artificial intelligence", *Machine Intelligence*, vol.4, pp.463-502, 1969.
- [6] Miki Ueno, Naoki Mori, Toshinori Suenaga, and Hitoshi Isahara. "Estimation of structure of four-scene comics by convolutional neural networks". In *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding*, MANPU@ICPR 2016, pp. 9:1–9:6, 2016.
- [7] 松下 光範, 山西 良典, 松井 勇佑, 岩田 基, 上野 未貴, 西原 陽子, 中村 聡史. "私のブックマーク「コミック工学」". 人工知能, Vol. 32, 2017.
- [8] Miki Ueno, Naoki Mori, and Keinosuke Matsumoto. "2-scene comic creating system based on the distribution of picture state transition". *Advances in Intelligent Systems and Computing*, Vol. 290, pp. 459–467. Springer, 2014.
- [9] 藤野 紗耶, 安井 瑛彦, 上野 未貴, 森 直樹, 松本 敬之亮. "Pictgent を用いた対話型絵本システムの提案", 第 60 回システム制御情報学会研究発表講演会, 315-4, 2016.
- [10] Yusuke Matsui, Kota Ito, Yuji Aramaki, Toshihiko Yamasaki, and Kiyoharu Aizawa. "Sketch-based manga retrieval using manga109 dataset". *CoRR*, Vol. abs/1510.04389, 2015.
- [11] Clément Guérin, Christophe Rigaud, Antoine Mercier, Farid Ammar-Boudjelal, Karell Bertet, Alain Bouju, Jean-Christophe Burie, Georges Louis, Jean-Marc Ogier, and Arnaud Revel. "ebdtheque: a representative database of comics". In *Proceedings of the 12th International Conference on Document Analysis and Recognition (ICDAR)*, 2013.
- [12] Neil Cohn, "You're a good structure, Charlie Brown: The distribution of narrative categories in comic strips", *Cognitive Science*, Vol. 38, No. 7, pp. 1317-1359, 2014.