

構造的組み合わせによる事象融合の計算モデル

A Computational Model of Event Blending by Structural Combination

秋元 泰介

AKIMOTO Taisuke

九州工業大学

Kyushu Institute of Technology

akimoto@ai.kyutech.ac.jp

Abstract

This study, with an aim of computational implementation of story creativity in a cognitive system, proposes a computational model of event blending based on the conceptual blending theory proposed by Fauconnier and Turner. The proposed model composes the representation of a new event via structural combinations of two input events, based on the mapping between input events. Furthermore, this study conducted an experiment to compare the output of the implemented model with human-generated ideas. The result demonstrated the model's potential of producing interesting ideas from given events.

Keywords — **Conceptual Blending, Story, Event, Memory, Computational Creativity, Cognitive System**

1. はじめに

物語の認知の仕組みを明らかにすることは、人間のよう人工知能を実現するための重要な課題の一つである。古くは Schank らにより、物語的な記憶構造に基づく人工知能・認知科学の研究が行われてきた[1, 2, 3]。Winston らの研究グループにおいても、物語を操る能力が人間の知能を特徴付ける要素であるという仮説[4]のもとに、物語の理解を中心に据えた人工知能研究が行われてきた。日本の人工知能研究においては、例えば岡田[5]や小方[6]の研究が挙げられる。

筆者は、特に、ストーリーを生み出す力が、人間的な知性ないし心を作るための基盤になると考える。これまで、個体(エージェント)が世界を表象する形として、「ストーリー」という概念を提案してきた[7, 8]。ここでストーリーと呼ぶのは、個体間のコミュニケーションを媒介する物語表現ではなく、物語的な構造を持つ内的表象である。物語的な構造というのは、端的に言うと複数の事象の並びのことであり、事象に関与する物や場所もそこに含まれる¹。ストーリーは、過去の経験の記憶、現在進行中の出来事、想像による未来や虚構、仮想、物語によるコミュニケーションを通して得られ

る情報などに共通する形として仮定される。そうすると、ストーリーを生成しながら物理的・社会的な環境と関わり合うこと、あるいは環境と関わり合いながらストーリーを生成することが、人間的な人工知能の基盤になる。

以上の考えのもとに、ストーリー生成を中心に据えた(人工)認知システムを作ること、またそれを通してストーリーを生み出す心の仕組みを探求すること、これが本研究の大局的な目的である。

人工知能研究において、ストーリーの創造性は様々な観点から研究されてきた。生成における中心的な処理に注目して分類すると、主に次のようなアプローチが挙げられる。

- プランニングによる登場人物の目標指向的な行動系列の生成[e.g., 10, 11].
- 物語構造のスキーマ的な形式化[e.g., 12, 13].
- 事例ベース推論や類推による生成[e.g., 14, 15].
- 物語の構造や生成過程の物語論的なモデル化ないしシステム化[e.g., 6, 16, 17].
- 統計的な機械学習やニューラル・ネットワークに基づくデータ駆動的な方法[e.g., 18, 19].

ストーリー生成は複合的な問題であるため、これらはそれぞれがその部分問題をなすという見方もできる。実際、複数の方法や考え方を組み合わせている研究も多い。しかし、具体的なタスクや機能(何らかの目的に沿ってストーリーを生成・出力すること)を中心として、様々な方法を組み合わせてシステムを構成するというやり方では、総合的な認知システムの基盤にはならないだろうし、ストーリーを生み出す心の仕組みを全体的に捉えることもできないだろう。

まず、ストーリーの生成を、心の具体的な機能や性質から切り離して、生成の原理的な仕組みを主題化する必要がある。ストーリー生成は、例えば、過去の経験の

¹ このような意味でのストーリーを認知システムの内部で如何に表現するかは難しい問題である。筆者は、マンガのような形で記号的かつマルチモーダルにストーリー＝世界を表現

するのが理想的であると考えている[9]が、本稿で示すモデルでは簡易的に、述語項構造に相当するシンボリック表現と単語の分散表現を組み合わせた形で事象を表す。

想起, 未来の予想や計画, 他者とのコミュニケーション, 他者の心的状態の想像, 新しい道具や環境の考案等, 心の様々な性質や機能の基盤になると考えられる. このような機能的な多面性は, 認知システムが作動する文脈や環境との関係のもとに生じると仮定しておき, その根底をなす原理に焦点を合わせるべきであろう.

人工知能研究においては, 具体的なタスクや問題を定め, 目標指向的なプロセスとしてそれを定式化するというのが一般的である. ストーリー生成の場合は, ストーリー内のキャラクターの目標[10, 11]や, ストーリー外の (仮想的な) 作者や語り手の目標に相当するもの[6, 14, 20]を, 生成の主要な駆動力としている研究が多く見受けられる. しかし, 目標指向的な見方は, ストーリーの生成を通して心の様々な機能や性質が生じるという筆者の考えには馴染みにくい. また, 目標が如何に生じるのかという問題もある.

筆者は, ストーリーの生成を, ストーリーあるいは多数のストーリーを含む記憶システムの自己組織的ないし創発的な動きとして捉えるべきであると考え. 目標に相当する要素は, ストーリーに内在するものとして, ストーリーの生成を通して生じる場合もあるという程度に考えておく (それが内側からストーリーの生成を方向付けたり, 他のストーリーの生成に影響を与えたりすることもあるだろう). ストーリーの生成を自己組織的な動きとして捉えることは, ストーリーが複雑な構造 (例えば部分間や部分-全体間の相互依存性や, 他のストーリーとの関係) を持つことや, ストーリーの生成が認知システムの総合的な働き (知覚-行動, 記憶, 概念的な認知等) を要するという点にも馴染みやすい.

以上の立場から, 当面の研究の道筋の一つは, 認知システムを構成する記憶システムの内部における, ストーリーの原始的で小さな動きを, 計算のレベルで理論化していくことである. ここでストーリーの原始的な動きとして想定されるのは, 例えば, 結合, 切断, 変形, 抽象, 関連付け, 一般化, 類似, 類比, 類推, 融合などである. 実際は, それぞれが複合的かつ相互に絡み合うような処理になると考えられるため, それらを徐々に総合していくことも重要である.

本稿では特に, 「融合」に焦点を合わせる. 筆者はこれまで, Fauconnier と Turner [21]による Conceptual Blending 理論及びその計算モデル化に関する研究に示唆を得て, ストーリー融合の全体的な構想を示した[22]. 本稿では, それを具体化していく過程における試みの一つとして, 二つの事象を構造的に組み合わせる新たな

事象を生成する計算モデルを提案する.

以降の構成は次のようになっている. 2 節では, Conceptual Blending 理論及びその計算モデル化に関する研究を概観する. 3 節では事象融合の仕組みを定式化し, 4 節でその実装を示す. 5 節では, 提案モデルの創造性を, 人による事象融合との比較を通して検証する. 6 節でまとめと今後の展望を述べる.

2. Conceptual Blending 理論

Conceptual Blending 理論[21]の概要を, 本稿で示すモデルに関わる部分を中心に説明する. この理論の計算モデル化に関する研究も概観する.

2.1 基本図式

Conceptual Blending 理論は, 人の創造的かつ日常的な思考の仕組みを, 複数の心的空間 (情報の小さなまとまり) の概念的な融合として説明する. その基本的な図式は図 1 のように表される.

元になる心的空間を入力空間 (input space) と呼び, それらの融合により生じる心的空間を融合空間 (blended space) または単に blend と呼ぶ. また, 二つ (複数) の入力空間の間の共通構造に相当するものが包括的空間 (generic space) に反映される. 包括的空間を通して, 入力空間の間に, 類似や同一性に基づく対応関係 (cross-space mapping ないし counterpart connections) が構成される. これが融合における土台となる. また, blend は, 入力空間から直接的に取り込まれる情報だけではなく, 創発的構造 (emergent structure) を含む場合がある. これは, 例えば, 複数の入力空間を組み合わせる際に生じる不調和を解消して, 新しい意味的まとまりを構成するような動きを通して生じるとされる.

Conceptual Blending は, 入力空間の間に, 類似や同一性に基づく対応関係を構成することが生成の中心になるという点で, 類推[23, 24]に似ている. しかし, 類推

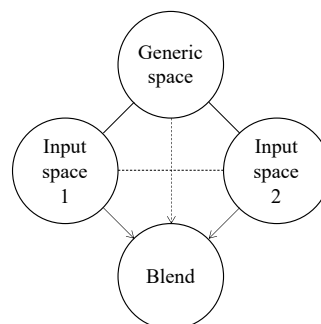


図 1 Conceptual Blending の基本図式 ([21]に基づく)

が一般的に元領域 (base) から対象領域 (target) への転移という図式で説明されるのに対して, Conceptual Blending は複数の入力空間からそれらとは異なる何かを生み出すような図式になっている。

2.2 統合ネットワーク (Integration Network)

Fauconnier と Turner [21]は, Conceptual Blending の仕組みを様々な観点から述べているが, ここではその一つとして, 概念空間間の関係を表す 4 種類の統合ネットワーク (simplex, mirror, single-scope, double-scope) を説明する。以下では入力空間が二つの場合を想定する。

Simplex ネットワークは, 抽象的なフレームと具体的な要素を融合する。例えば, 「家族」フレームと, 複数の人 (例: 「タロウ」と「アコ」) の融合により, 「タロウはアコの父」のような意味が構成される。

Mirror ネットワークは, 共通の組織化フレームにより構成される二つの入力空間を融合する。例えば, “Great American II [in 1993] is 4.5 days ahead of Northern Light [in 1853]”という表現の概念構造 (blend) は, 異なる時代における 2 隻の船それぞれの航海を融合して, 新たに「競走」という意味が付加されたものとされる。

Single-scope ネットワークは, 異なる組織化フレームによる二つの入力空間を融合するが, その際, 何れか一方の入力空間のフレームが blend においても使用される。例えば, “Murdoch [CEO1] knocks out Iacocca [CEO2]”という比喩的な表現の概念構造 (blend) は, 「ビジネス」の状況を「ボクシング」のフレームに当てはめて再構成したものとされる。

Double-scope ネットワークは, 異なる組織化フレームによる二つの入力空間を, 新しい組織化フレームを構成することによって融合する。例えば, 「デスクトップ・インタフェース」(blend) は, 「事務作業」と「計算機コマンド」を, これらを組み合わせた新しいフレームにより融合したものとされる。

2.3 Conceptual Blending の計算モデル

Conceptual Blending 理論を計算のレベルへと精緻化する研究や, それをもとに創造的な計算システムを作る試みも行われている。例えば, Goguen と Harrell [25] は, 詩的な物語を生成するシステムの一部として, 二つの概念を融合する仕組みを形式化している。例として, 「house」(家) の概念構造 (“on land”, “live-in resident”) と「boat」(船) の概念構造 (“on water”, “ride passenger”) を組み合わせて, 「houseboat」という新しい概念構造を

生成する問題が示されている。Confalonieri らの編著[26]では, Conceptual Blending の計算モデルに関するより体系的な研究プロジェクトが示されている。

Conceptual Blending の計算モデル化は様々な要素課題に分けることができる。まず, 入力空間 (多くの場合は houseboat のような概念構造) の間に, 適切な包括的空間を構成することが一つの問題とされる。Eppe ら [27]の研究では, 事例ベース推論の amalgam 理論[28]に基づいて, 概念要素を適切なレベルまで一般化する仕組みを構築している。Besold [26: ch.5]は, 類比を通じた一般化について検討している。Hedblom ら[29]は, 概念の非言語的な側面として, イメージ・スキーマに基づく包括的空間の形成を論じている。

Conceptual Blending の計算においては, 一般的に, 概念間の組み合わせ方に多数の可能性があり, その中から「良い」融合を選別するための基準を形式化することも一つの問題となる。例えば Eppe ら[27]の研究では, Conceptual Blending 理論における最適性基準 (optimality principles) をもとに, blend における情報の量やバランス等の観点からそれを形式化している。一方, 問題領域に特化した価値基準を導入する研究もある[26: ch.2]。

上に挙げた研究の多くは, 一般的な概念 (例えば「家」と「船」) の融合を問題としている。このような処理もストーリーの創造性の基盤になると考えられるが, 本研究は, ストーリーの具体的な内容を融合することに焦点を合わせる。その初歩的な試みとして, 今回はストーリーの基本的な要素となる「事象」を融合する計算モデルを提案する。現状では, 前述した既存の研究で示されている方法を参考にしている面も大きい。しかし, 元の Conceptual Blending 理論[21]が示している例の多くは, 具体的な状況やストーリーに相当する概念構造を扱っているため, ストーリーの融合を問題とすることは, Conceptual Blending の精緻化という点でも有意義であろう。

3. 事象融合の定式化

本稿で提案するモデルでは, 事象融合を, 二つの入力事象 E_L と E_R を構造的に組み合わせて融合事象 E_B を構成する処理として定式化する。

3.1 事象の表現

事象は, 述語項構造ないし格フレーム[30]に相当する形で表現する。すなわち, 動詞 (的概念) と一つ以上の

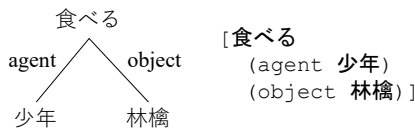


図2 事象の構造

項(名詞的概念)を, 意味役割を明記して結合する形で, 事象を表す. 例えば, 「少年が林檎を食べる」という意味の事象は図2のようになる.

以下では, ある事象 E_x の動詞部を v_x , i 番目の項を e_i^x と表記する. 項 e_i^x の意味役割(agent等)を指す場合は $r[e_i^x]$, 値(名詞的概念)を指す場合は $n[e_i^x]$ と表記する.

3.2 融合事象の基本要件

本研究はストーリーの創造性の根源的な原理に焦点を合わせるので, 何がストーリーの生成を方向付けるかについても, 具体的な目標や用途ではなく, 抽象的な水準で考える. Eppeら[27]のように, Conceptual Blendingの最適性基準[21]を参考にするという手段もあるが, まずはより広い観点から検討したい.

Boden [31]は, 創造性を端的に“the ability to generate novel, and valuable, ideas”と定義している. この定義にあるように, 「新しさ」と「価値」が創造の基本的な条件になると考えられる. 新しさについては, Boden[31]の用語を使うと, 個人にとっての新しさ (psychological) と, 社会における歴史的な新しさ (historical) に分けられる. 一方, ストーリーの価値には, ストーリーを実行したり他者に伝えたりすることによって生じる社会的な価値と, 個体自身にとっての価値の両面がある. 例えば, 密林で暮らす中で, 新しくて効率的な狩猟方法を考案することは, 当の個体にとって有用であり, それを他者と共有すれば社会的な価値にもなる.

しかし, 新しさと価値は, 何れも時間的な文脈(社会的な時代背景や個体の発達過程)や, 用途に依存する面が大きい. そこで本研究では, これらをより原始的な水準へと抽象化し, 次の2つの要件を定める.

- A) 差異: 入力事象に対する差異を生み出すことが「新しさ」の必要条件である.
- B) 意味: 生成された事象が意味をなすことが「価値」の必要条件である.

これらをモデルの設計における基本的な指針とする.

なお, ここで「意味をなす」というのは, 現実世界において成立するか否かということではない. 例えば, 「食べる (agent 林檎) (object 少年)」という事

象は, 現実世界では起こり得ないが, 空想的な世界においては成立し得る. 一方, 「食べる (agent 林檎) (from 月)」のような事象は意味をなしにくいだろう.

3.3 包括的空間と事象間マッピング

事象融合における主要な処理の一つは, 包括的空間を通して, 入力事象間に対応関係(マッピング)を構成することである.

Conceptual Blending 理論[21]における包括的空間は, 入力空間間を包括するカテゴリ的な語やフレーズにより表されている場合が多い(実際のところ, この理論には計算レベルで見ると曖昧な点が多くある). これを表面的に真似て, WordNetのような階層的語彙体系(シソーラス)を利用するという手段も考えられるが, 筆者は, 広い意味での類似に基づいて入力事象間に対応関係を構成することが包括的空間の基本的な役割であると解釈する.

提案モデルでは, 事象を構成する要素(概念)間の類似度を計算する仕組みの一つとして, 単語の分散表現モデルの一つである word2vec の skip-gram [32]をここに位置付ける. 加えて, 事象の述語項構造を構成する関係(意味役割)もまた, 要素間の類似を構成する要因の一つと見なせる. 以上の考えに基づいて, 入力事象間のマッピングを以下のように定義する.

まず, 事象間マッピングを, 動詞と項それぞれの1対1対応を表す集合とし,

$$M_{L,R} = \{m_1, \dots, m_n\} \quad (1)$$

と表す.

ここで, 二つの事象それぞれの動詞部は常に対になると仮定する. 項の対応関係は, 以下の類似度計算式に基づいて決める.

$$\text{sim}(e_i^L, e_j^R) = \text{match}(r[e_i^L], r[e_j^R])\alpha + \cos(n[e_i^L], n[e_j^R])(1 - \alpha) \quad (2)$$

$\text{match}()$ は, 二つの項の意味役割が一致する場合1, 不一致の場合0を取る. $\cos()$ は, 項の値に対応する単語ベクトル間のコサイン類似度である. α はこれらの比重を決める係数であり, $0 \leq \alpha \leq 1$ とする.

事象間マッピングにおいては, 上記の類似度がより高い項どうしを優先的に対応付け, 類似度が閾値 β 以下の項間は対にしないこととする. 対応する項が無い場合は空記号 ε と対にする.

例えば, E_L が「ノックアウトする (agent ボクサー

(1) (counter-agent ボクサー(2)), E_R が[爆破する (agent 兵隊) (object ビル)]の場合の可能なマッピング $M_{L,R}$ の一つは, {(ノックアウトする, 爆破する), (agent:ボクサー(1), agent:兵隊), (counter-agent:ボクサー(2), ε), (ε , object:ビル)}となる. 重み係数 α や閾値 β の値によって, 「ボクサー(2)」と「ビル」が対になる可能性もある.

3.4 事象融合のモード

本モデルでは, 2.2節で説明した4種類の統合ネットワークの中の, single-scope と double-scope に対応する融合方法を扱う. それぞれ SS 及び DS モードと呼ぶ.

以下に示される方法は, 体系的な検討に基づくものではなく, single-scope と double-scope の各融合を具現化する方法を模索する過程で出来たものである. 何れも数ある可能性の中の一つに過ぎない.

3.4.1 SS モード

SS モードでは, 二つの入力事象に主従関係が付与される². 以下ではこれらを主事象, 従事象と呼び分ける. 図3にこの構造を例示する. このモードは, まず, 主事象の述語項構造を融合事象に反映する. そして, 事象間マッピングを通して, 従事象の要素の一部を, 主事象の要素と対にする形で融合事象に取り込む. その後, 対になった単語を意味的に合成する. 動詞部は, “ v_a as-if-to v_b ” (v_b するかのよう v_a する) という直喩の形で合成する. ここで v_a は主事象の動詞, v_b は従事象の動詞を表す. 項対の合成方法は DS モードと共通化されているため後の3.5節で説明する.

3.4.2 DS モード

DS モードでは, 二つの入力事象をほぼ同等に扱い, 二つの述語項構造を意味的に重ね合わせる形で組み合わせる. 図4にこの構造を例示する.

動詞部は“ v_1 by v_2 ” (v_2 することによって v_1 する) という形で合成される. このように合成すると, 基本的に, v_2 が v_1 を行う「手段」となる. 動詞の並べ方には, “ v_L by v_R ” (LR形式) と“ v_R by v_L ” (RL形式) の2通りがある.

融合事象の述語項構造は, 二つの述語項構造を構造的に重ね合わせることによって新たに構成される. 事象間マッピング ($M_{L,R}$) において, 対になった項の意味役割に相違がある場合は, v_1 側の入力事象における意

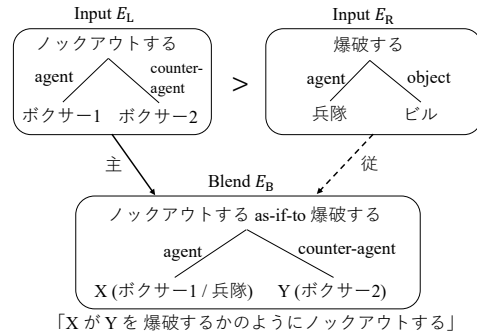


図3 SSモードによる融合の例

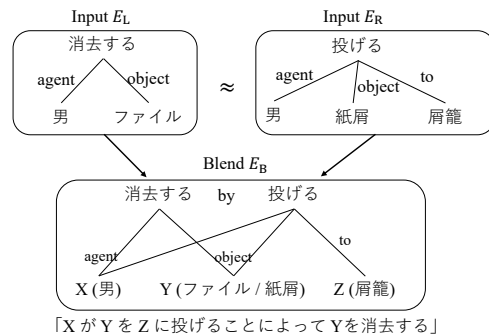


図4 DSモードによる融合の例.

味役割が優先される.

図4の融合事象は, $M_{L,R} = \{(\text{消去する}, \text{投げる}), (\text{agent:男}, \text{agent:男}), (\text{object:ファイル}, \text{object:紙屑}), (\varepsilon, \text{to:屑籠})\}$ の時に構成される構造である. この融合事象の述語項構造は, [消去する by 投げる (agent[1, 2] X) (object[1, 2] Y) (to[2] Z)]と表される (項の値は後述する合成処理により決まる). 意味役割の後の角括弧内の番号は, それがかかる動詞 (v_1 と v_2)を表す.

3.5 対になった項の値の合成

3.4節で述べた融合処理によって対になった項の値 (単語) を合成する操作を3種類定義する. 各操作は, 単語対から可能な値の集合を返す.

最初の操作は, 二つの単語 w_1 と w_2 の単純な組み合わせである. これは次のように定義される.

$$\sigma_c(w_1, w_2) = \begin{cases} \{(w_1), (w_2), (w_1, w_2)\}, & (w_1 \neq w_2) \\ \{(w_1)\}, & (w_1 = w_2) \end{cases} \quad (3)$$

融合事象において対になった項の値には, まずこの操作が適用されて, 一方を取る場合と, 両方を取る場合に

うな形になっている. この部分は今後の研究で修正・拡張することになるだろう.

² 実はこの解釈には少し不十分な点がある. 2.2節で挙げた「ボクシング+ビジネス」の例は, 単純な主従関係ではなく, 一方の構造 (フレーム) に, 他方の内容を当てはめるよ

分岐する。一方を取る場合は、それが融合事象における項の値となる。両方の値を取る場合は、二つの単語を直喩 (σ_{simi}) または複合語 (σ_{comp}) の形で意味的に合成する。これらの操作は次のように定義される。

$$\sigma_{simi}(w_1, w_2) = \{(w_1 \text{ like } w_2), (w_2 \text{ like } w_1)\} \quad (4)$$

$$\sigma_{comp}(w_1, w_2) = \{(w_1 w_2), (w_2 w_1)\} \quad (5)$$

全体として、二つの異なる単語の対には 6 パターンの合成がある。例えば、図 3 の agent 部分の可能な値は、{(ボクサー), (兵隊), (ボクサー like 兵隊), (兵隊 like ボクサー), (兵隊ボクサー), (ボクサー兵隊)}である。

4. 実装

以上の定式化に基づいて、事象融合の計算モデルを構築した。実装言語は Scala である。このモデルは、外部から与えられた二つの入力事象をもとに一つの融合事象を生成する。これらの事象は前述した述語項構造の形式で表されるが、テンプレートを用いた簡易的な文生成処理により、融合事象の文表現も出力される。

モデルは以下の設定項目に従って作動する。

- 融合モード: {SS, DS}
- SS モードにおける主事象: {L, R}
- DS モードにおける動詞順序: {LR, RL}
- 項の意味的合成方法: $\{\sigma_{simi}, \sigma_{comp}\}$
- 事象間マッピングの重み係数 α 及び閾値 β の値。

これらを決めることで解空間が限定される。そして、入力事象に対する差異を基準とする選択を経て、最終的に一つの融合事象が生成される。

なお、本モデルは将来的な拡張を見据えて設計されているため、現段階では概念的または仕組み的に冗長に見える部分もあるだろう。

4.1 構造的図式

図 5 は、本モデルにおける事象融合を構造的に表した図式である。 E_L と E_R は外部から与えられる入力事象である。本モデルでは、包括的空間に相当する構造は明示的に表現せずに、それを通して形成される事象間マッピング $M_{L,R}$ をもとに事象を融合する。 $C_{L,R}$ は、融合の過程で作られる可能な組み合わせ構造の集合である。その中の一つを選択して最終的な融合事象 E_B が構成される。この選択は、入力事象に対する差異に基づいて行われる。この差異を計算する際にも、入力事象と組み合わせ構造の間にマッピング $M_{c_i,L}$ 及び $M_{c_i,R}$ が構成される。

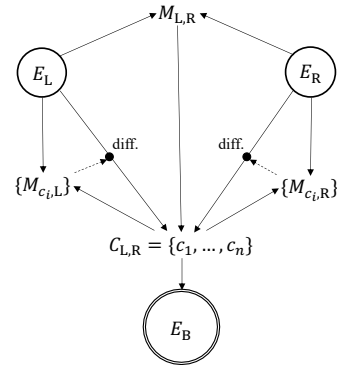


図 5 事象融合の構造的な図式

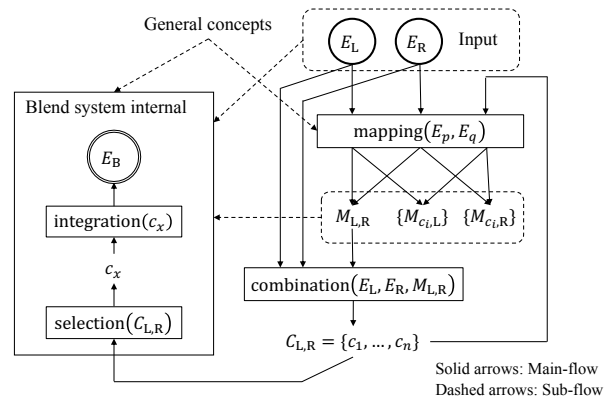


図 6 事象融合の処理の流れ

4.2 処理の流れ

事象融合の処理の流れは図 6 のようになっている。以下、手続きに沿って説明する。

- (1) $mapping(E_L, E_R)$: 入力事象間にマッピング $M_{L,R}$ を構成する。
- (2) $combination(E_L, E_R, M_{L,R})$: 入力事象間のマッピングに基づいて、可能な組み合わせ構造の集合 $C_{L,R}$ を構成する。この段階では、SS モードまたは DS モードによる述語項構造の融合 (3.4 節) と、対になった項に対する単純選択操作 (σ_c) が適用される。
- (3) $selection(C_{L,R})$: 組み合わせ構造の集合において、4.3 節で述べる基準により、入力事象対に対する差異が大きいものを一つ選択する。
- (4) $integration(c_x)$: 選択された構造の項部分で対になった単語を、3.5 節で述べた 2 種類の意味的合成方法、すなわち直喩 (σ_{simi}) または複合語 (σ_{comp}) の何れかの形で合成する。この結果が融合事象 E_B となる。

図 6 中の general concepts (一般概念) は、skip-gram [32]により構築された単語のベクトル表現であり、mapping 処理における事象要素間の類似度計算 (式 2) や、後述する差異の計算において使用される。

4.3 組み合わせ構造の選択基準

複数の組み合わせ構造の中からの選択 (selection) は、入力事象に対する差異に基づいて行われる。この計算方法を以下に説明する。

まず、組み合わせ構造 c_i と一方の入力事象 E_x の間の差異を、事象間マッピング $M_{c_i,x}$ における各対の非類似度の平均とし、次のように定義する。

$$\text{diff}(c_i, E_x) = \frac{1}{|M_{c_i,x}|} \sum_j \left(1 - \cos \left(n[m_j^{c_i}], n[m_j^x] \right) \mu_j \right) \quad (6)$$

$n[m_j^{c_i}]$ と $n[m_j^x]$ はそれぞれ、 $M_{c_i,x}$ における j 番目の対における項の値または動詞を表し、 $\cos()$ はこれらの単語間のコサインである。但し、一方が空 (ϵ) の対は計算の対象外とする。 μ_j については以下で説明する。

式6において、 c_i 側の値 ($n[m_j^{c_i}]$) が二つの単語の組み合わせの場合は、特別に次のような処理がなされる。まず、二つの単語ベクトルの重心をその値のベクトルと見なす。しかし、二つの単語の重心ベクトルは、元の二つの単語両方との類似度が高くなる傾向がある (例えば、「兵隊+ボクサー」は、「兵隊」と「ボクサー」両方に似ている)。そのため、重心ベクトルによる類似度をそのまま用いると、単語の組み合わせにより生じる意味的な変化が差異の値に反映されない。そこで、似ていない単語が組み合わせられる程、新たな意味が生じて入力事象に対する差異が大きくなると仮定し、式6における各対の類似度 (\cos) にかかる係数 μ_j を次のように定義する。

$$\mu_j = \begin{cases} 1, & \left(n[m_j^{c_i}] = 1 \right) \\ \cos \left(n_1[m_j^{c_i}], n_2[m_j^{c_i}] \right), & \left(n[m_j^{c_i}] = 2 \right) \end{cases} \quad (7)$$

この値が小さい程、差異の値が大きくなる。

selection 処理では、 $\text{diff}(c_i, E_L)$ と $\text{diff}(c_i, E_R)$ の合計が上位 20%に入るものの中から、乱数により一つが選択される。この 20%という値に深い意味はないため、任意の値に変更してもよい。

5. 人による事象融合との比較

モデルの創造性の程度や性質を検証するための実験を行った。この実験は、何らかの仮説を検証することではなく、人との比較を通して相対的にモデルの性質を分析することを意図したものである。

表1 実験用の入力事象対

	E_L	E_R
P1	学生が教授を論破する [論破する (agent 学生) (counter-agent 教授)]	ボクサー1がボクサー2をノックアウトする [ノックアウト (agent ボクサー(1)) (counter-agent ボクサー(2))]
P2	男が女にキスする [キスする (agent 男) (counter-agent 女)]	兵隊がビルを爆破する [爆破する (agent 兵隊) (object ビル)]
P3	未亡人が外国に旅立つ [旅立つ (agent 未亡人) (to 外国)]	栗鼠 (りす) が団栗 (どんぐり) を探す [探す (agent 栗鼠) (object 団栗)]
P4	老人がファイルを消去する [消去 (agent 老人) (object ファイル)]	秘書が紙屑 (かみくず) を屑籠 (くずかご) に投げる [投げる (agent 秘書) (object 紙屑) (to 屑籠)]
P5	警官がパトカーを運転する [運転 (agent 警官) (object パトカー)]	白熊が海豹 (あざらし) を追う [追う (agent 白熊) (counter-agent 海豹)]
P6	忍者が城に忍び込む [忍び込む (agent 忍者) (to 城)]	医者が国王を診察する [診察 (agent 医者) (counter-agent 国王)]

5.1 方法

表1に示す6対の入力事象を用意した。作成者は筆者であり、本モデルの仕組みを考慮して作られている点には留意が必要である。これらの事象対を用いて、計算モデルによる融合事象と、同様の状況で人により作成された融合事象を比較するアンケート調査を行った。

5.1.1 モデルによる融合

計算モデルに表1に示した6つの入力事象対を与えて、それぞれから2つ、合計12個の融合事象を用意した。その一覧を表2の融合事象C欄に示す。入力事象対P1-3にはSSモード、P4-6にはDSモードを適用する。各入力事象対について、SSモードにおいては、主事象が E_L の場合と E_R の場合の2通り、DSモードにおいては動詞の並びがLR形式の場合とRL形式の場合の2通りの出力を用意した。それぞれについて、6回生成を行い、その結果の中から筆者が一つを選抜した。なお、6回のうちの3回は項の意味的合成に直喩 (σ_{simi}) を用い、残りの3回は複合語 (σ_{comp}) を用いた。事象間マッピングにおけるパラメータは、経験的に、入力事象間 ($M_{L,R}$) においては $\alpha = 0.2, \beta = 0.2$ 、入力事象と組み合わせ構造の間 ($M_{c_i,L}$ と $M_{c_i,R}$) においては $\alpha = 0.9, \beta = 0.0$ とした。

5.1.2 人による融合

人による融合事象は、情報工学系の大学生4名により作成された。各作成者には、表1に示した6対の入力事象 (文) が示されたカードが渡され、それぞれにつ

いて、「二つの事象文の“意味”をクリエイティブに融合して新しい事象文(1文)を創ること」及び「なるべく平易な文体で書くこと」が指示される。一人あたりの所要時間は概ね15分であった。各事象対について合計4つの融合事象が得られるため、その中から筆者が面白いと思うもの2つを選抜し、前述した計算モデルによる12個の融合事象と組にした。人による融合事象は表2の融合事象H欄に示されている。

5.1.3 アンケート

以上の方法で作成された12組の比較データを用いて、アンケート調査を行った。回答者には、元になっている入力事象対と、比較対象である二つの融合事象文が「2つの事象の意味を混ぜ合わせて新しい事象を作るという課題の答え」であることが伝えられる。この時、作成者に関する情報は示されない。コンピュータによる出力があることも明示されないが、実験の趣旨として、人工知能研究に関する調査であることは伝えられる。

回答者は、各組について次の3つの質問に回答する。

- Q1. どちらがよりクリエイティブですか？(一方または引分を選択)
- Q2. それぞれについて、どのくらいクリエイティブかを7段階で採点してください。(1~7点)
- Q3. それぞれについて、文の意味が通りますか？(はい/いいえ)

Q1とQ2は創造性の程度を総合的に測るための質問であり、「クリエイティブ」をどのように解釈するかは個々の回答者に委ねる。Q3は、3.2節で述べた「意味」の要件に対応する質問である。

回答の制限時間は15分とした。融合事象の提示順序が回答に影響を与える可能性があると考え、約半数の回答者にはNo.1から12という順序で、残る半数にはその逆順で提示した。

5.2 結果と考察

上記アンケートの回答を45名の大学生(情報工学部、18~23歳の男女)から集めた。各質問項目の結果を表2にまとめる。全体として、人による融合事象が、何れの質問項目においても、モデルを上回っている。しかしながら、Q1(勝数)やQ2(創造性の点数)においてはモデルと人の間に大きな差は見られない。また、いくつかの組(No.2,4,6,7)においては、Q1及びQ2の値においてモデルが人を上回っている。

モデルが人を上回った組を見ると、モデルの出力が他の組よりも面白かったというよりも、人側の融合事

象が簡素である場合が多いようである。人工知能の立場から見ると、例えば、表1のP3から「未亡人が仕事を探す」(表2, No.6)という事象を作ることにも、かなり高度な認知が介在しているように思われるが、人の評価はあまり高くなかったようである。

一方、文の意味が通るかどうかを問うQ3については、モデル側、特にDSモード(No.7-12)において低い値となっている。DSモードは比較的複雑な構造を生成するため、簡易的な文生成処理がその意味を適切に表現できていないという問題もあるだろう。しかし、意味が通らないことが、必ずしも創造性の勝数(Q1)や点数(Q2)を下げる要因にはなっていないようである。人は、不手際のあるアイデアにも面白さや奇抜さを見出すことができるということだろう。また、奇抜なアイデアの意味が解釈困難であることも珍しくない。筆者としては、表2の融合事象の中に、全く意味を成さないものはないと考えている。

しかし、ここで一つ問題になるのは、今回の実験状況において、人が融合事象文を解釈すること自体に、創造的ないし想像的な認知が介在するという点である。例えば、モデルによる融合事象No.9(表2)は、「警官白熊が、海豹を追うことによって、パトカーを運転する」である。これは、白熊(の警官)が普通に車を運転するのではなく、海豹を追うような行為をすることによって車が走る、という状況を想像すれば意味が成り立つ。車中で白熊が海豹を追うという状況は想像しにくいかも知れないが、どこか広い所で白熊が海豹を追うと車が遠隔操作されるとか、車中の白熊がVR空間において海豹を追うと車が走る、などという状況は想像できる。人は、このように想像力を駆使して融合事象文の意味を解釈することができる。モデルによる融合事象No.8や11も、解釈を広げれば意味をなすだろう。しかし、事象融合は認知システムの内的な動きとして位置付けられるため、上記のような意味も含めて、まとまりのある事象ないしストーリーが生成されるべきである。Conceptual Blending理論[21]においては、融合を通して新たに生じる構造のことを創発的構造と呼んでいる。上の例はやや極端かも知れないが、このような創発的な生成の仕組みについては、今後さらなる検討が必要である。

次に、計算モデルと人の間の質的な違いに着目する。計算モデルによる融合方法は、大きく分けて2種類(SSとDS)に決まっているため、複数の出力を並べると、形式的に単調な印象を受けるだろう。一方、人はより多

表2 計算モデルによる融合事象と人による融合事象の比較結果

No. [入力]	融合事象 (文)		Q1. 勝数: C-H [引分]	Q2. 平均点 [SD]		Q3. “はい”の 割合 (%)	
	C: 計算モデル [動作設定]	H: 人 [作成者 ID]		C	H	C	H
1 [P1]	学生が、ボクサーのような教授を、ノックアウトするかのように論破する。[by SS(L, Simi)]	教授は学生に言葉という暴力でノックアウトされた。[by H1]	16-22 [7]	4.49 [1.56]	4.67 [1.63]	80.0	97.8
2 [P1]	学生ボクサーが、教授ボクサーを、論破するかのようにノックアウトする。[by SS(R, Simi)]	ボクサー1 が世界王者をノックアウトする。[by H3]	31-12 [1]	4.67 [1.79]	3.33 [1.88]	68.9	100.0
3 [P2]	兵隊のような男が、女に、爆破するようにキスする。[by SS(L, Simi)]	兵隊がキスしている男女を爆破する。[by H3]	16-21 [8]	4.84 [1.70]	5.04 [1.76]	79.5	97.8
4 [P2]	男兵隊が、キスするかのようにビルを爆破する。[by SS(R, Simi)]	兵隊がビルの中にいた女にキスすると、ビルが爆破した。[by H2]	24-15 [6]	5.30 [1.58]	4.64 [1.71]	72.7	97.8
5 [P3]	栗鼠未亡人が、外国へ旅立つ。[by SS(L, Simi)]	栗鼠の顔をした未亡人が団栗を探すために外国に旅立つ。[by H4]	7-25 [13]	3.44 [1.53]	4.62 [1.51]	55.6	93.3
6 [P3]	未亡人のような栗鼠が、団栗を、旅立つように探す。[by SS(R, Comp)]	未亡人が仕事を探す。[by H3]	27-14 [4]	4.20 [1.78]	3.11 [1.98]	48.9	100.0
7 [P4]	老人のような秘書が、紙屑のようなファイルを、屑籠へ投げることによって、消去する。[by DS(LR, Simi)]	執事がファイルを消去する。[by H3]	25-12 [7]	4.40 [1.57]	3.56 [1.76]	91.1	100.0
8 [P4]	秘書老人が、ファイル紙屑を、消去することによって、屑籠へ投げる。[by DS(RL, Comp)]	年老いた秘書がくずかごにファイルを投げる。[by H4]	16-19 [10]	3.80 [1.86]	3.82 [1.73]	15.6	95.3
9 [P5]	警官白熊が、海豹を追うことによって、パトカーを運転する。[by DS(LR, Comp)]	警官が海豹を運転する。[by H4]	9-22 [13]	4.05 [1.80]	4.89 [1.80]	28.9	55.6
10 [P5]	警官のような白熊が、パトカーを運転することによって、海豹を追う。[by DS(RL, Simi)]	白熊の警官がパトカーに乗って犯人の海豹を追う。[by H1]	11-23 [11]	4.20 [1.64]	5.07 [1.24]	71.1	93.2
11 [P6]	忍者のような医者が、国王のような城を診察することによって、(城に) 忍び込む。[by DS(LR, Simi)]	忍者が医者ふりをして、城に忍び込み国王を診察するふりをして暗殺した。[by H2]	14-28 [3]	3.98 [2.03]	4.91 [1.95]	20.0	100.0
12 [P6]	医者のような忍者が、城のような国王に、忍び込むことによって、(国王を) 診察する。[by DS(RL, Simi)]	城に忍び込んだ医者が忍者を診察する。[by H4]	19-19 [7]	4.27 [1.97]	4.22 [1.70]	40.0	92.9
Total			215-232 [90]	4.30	4.32	56.0	93.6

様な形で事象を混ぜ合わせているように見える。例えば、人による融合事象 No.5 (表2) は、「栗鼠の顔をした未亡人が団栗を探すために外国に旅立つ」である。この文からは、身体(栗鼠の顔をした未亡人)や目的(団栗を探すために旅立つ)の合成操作が読み取れる。このような事例に特化した融合方法を定式化することはできるかも知れないが、状況(入力事象対)に応じて柔軟に多様な構造を生み出すためには、与えられた入力事象だけを見て処理を行うという現在の枠組みでは不十分である。前述した創発的構造の問題にも通じるが、二つの事象の融合という比較的原始的な動きであっても、記憶や常識(に相当する何らかの仕組み)を含む認知システムの総合的な働きとして扱う必要がある。

6. まとめと展望

本稿では、二つの事象を構造的に組み合わせて新しい事象を構成する事象融合の計算モデルを提案した。

人による事象融合との比較を通して、提案モデルによる創造的な事象生成の可能性を多少は示すことができたものの、モデルに含まれる処理はまだ限定的である。今後取り組むべき主な問題を以下に挙げる。

SSモードの融合方法(比喩的な合成)は比較的一般的であると思われるが、DSモードの融合方法(述語項構造の手段的な重ね合わせ)はやや特殊かも知れない。新しい構造を作って融合するという問題は、入力事象間の直接的な組み合わせという枠を超えて、より総合的な認知プロセスとして考えるべきだろう。この問題は、5.2節で述べたように、創発的構造を生み出す仕組みにも関連する。その他、simplex及びmirrorネットワークに相当する融合方法も検討中である。

現在のモデルでは、入力事象対に対する差異を、生成の基本的な駆動力としている。この差異の計算をより質的なものにするのも課題の一つである。この点については、Conceptual Blendingの最適性基準[21, 27]が参

考になるかも知れない。

事象融合は認知システムの内部に位置付けられるため、入力事象対を外部から受け取るだけでなく、自らそれを形成することもできなければならない。これについては、多数のストーリー（記憶）の中から、創造的な融合が生じるような事象（ストーリー）どうしを自動的に関連付ける仕組みが主な問題となる。

事象融合からストーリー融合への拡張の道筋の一つとして、事象融合による局所的な処理がストーリーの全体構造に波及するような仕組みを考えていく予定である。これもまた創発的構造に関わる問題である。

本研究の主目的は、認知システムにおけるストーリー生成の原理を探求することであるが、Conceptual Blending 理論の計算モデル化としての側面もある。現時点では、Conceptual Blending の計算モデルとしての新奇性や貢献は十分に明確化されていないため、他の研究との理論的な関連付けや比較を通してそれを明確にしていくことも今後の課題の一つとしたい。

謝辞

本研究は JSPS 科研費 JP18K18344 の支援を受けた。

参考文献

- [1] Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Lawrence Erlbaum.
- [2] Schank, R. C. (1982). *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge University Press. (黒川 利明・黒川 容子 (訳)(1988). *ダイナミック・メモリ*. 近代科学社)
- [3] Schank, R. C. (1990). *Tell me a story: Narrative and intelligence*. Northwestern University Press. (長尾 確, 長尾 加寿恵 (訳) (1996). *人はなぜ話すのか*. 白揚社)
- [4] Winston, P. H. (2012). The right way. *Advances in Cognitive Systems*, 1, 23-36.
- [5] 岡田 直之 (2009). 脳に宿る心—認知科学・人工知能から神秘の世界に迫る. オーム社.
- [6] Ogata, T. (2016). Computational and cognitive approaches to narratology from the perspective of narrative generation. In T. Ogata, & T. Akimoto (Eds.), *Computational and Cognitive Approaches to Narratology* (pp. 1-74). IGI Global.
- [7] Akimoto, T. (2018). Stories as mental representations of an agent's subjective world: A structural overview. *Biologically Inspired Cognitive Architectures*, 25, 107-112.
- [8] Akimoto, T. (2019). Narrative structure in the mind: Translating Genette's narrative discourse theory into a cognitive system. *Cognitive Systems Research*, 58, 342-350.
- [9] 秋元 泰介 (2019). 世界をストーリーとして表象する Cogmic Space のコンセプト. 第 63 回ことば工学研究会資料 (pp. 83-101).
- [10] Meehan, J. R. (1980). *The metanovel: Writing stories by computer*. Garland Publishing.
- [11] Riedl, M. O., & Young, R. M. (2010). Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, 39, 217-267.
- [12] Pemberton, L. (1989). A modular approach to story generation. *Proc. 4th Conference on European Chapter of the Association for Computational Linguistics* (pp. 217-224).
- [13] Bringsjord, S., & Ferrucci, D. A. (1999). *Artificial intelligence and literary creativity: Inside the mind of BRUTUS, a storytelling machine*. Lawrence Erlbaum.
- [14] Turner, S. R. (1994). *The creative process: A computer model of storytelling and creativity*. Lawrence Erlbaum.
- [15] Ontañón, S., & Zhu, J. (2011). The SAM algorithm for analogy-based story generation. *Proc. 7th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* (pp. 67-72).
- [16] Gervás, P. (2016). Computational drafting of plot structures for Russian folk tales. *Cognitive Computation*, 8, 187-203.
- [17] 小方 孝 (2007). プロップから物語内容の修辞学へ—解体と再構成の修辞を中心として. *認知科学*, 14, 532-558
- [18] Fan, A., Lewis, M., & Dauphin, Y. (2018). Hierarchical neural story generation. *Proc. 56th Annual Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 889-898).
- [19] McIntyre, N., & Lapata, M. (2009). Learning to tell tales: A data-driven approach to story generation. *Proc. Joint Conference of the 47th ACL and IJCNLP* (pp. 217-225).
- [20] Pérez y Pérez, R., & Sharples, M. (2001). MEXICA: A computer model of a cognitive account of creative writing. *Journal of Experimental & Theoretical Artificial Intelligence*, 13, 119-139.
- [21] Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. Basic Books.
- [22] Akimoto, T. (2019). Theoretical framework for computational story blending: From a cognitive system perspective. *Proc. 10th International Conference on Computational Creativity* (pp. 49-56).
- [23] Gentner, D., & Forbus, K. D. (2011). Computational models of analogy. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 266-276.
- [24] Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. MIT Press. (鈴木 宏昭・河原 哲雄 (監訳) (1998). *アナロジーの力*. 新曜社)
- [25] Goguen, J. A., & Harrell, D. F. (2010). Style: A computational and conceptual blending-based approach. In S. Argamon, K. Burns, & S. Dubnov (Eds.), *The structure of style* (pp. 147-170). Springer-Verlag.
- [26] Confalonieri, R., Pease, A., Schorlemmer, M., Besold, T. R., Kutz, O., Maclean, E., & Kaliakatsos-Papakostas, M. (Eds.) (2018). *Concept invention: Foundations, implementation, social aspects and applications*. Springer.
- [27] Eppe, M., Maclean, E., Confalonieri, R., Kutz, O., Schorlemmer, M., Plaza, E., & Kühnberger, K. U. (2018). A computational framework for conceptual blending. *Artificial Intelligence*, 256, 105-129.
- [28] Ontañón, S., & Plaza, E. (2010). Amalgams: A formal approach for combining multiple case solutions. *LNCIS 6176* (pp. 257-271). Springer.
- [29] Hedblom, M. M., Kutz, O., & Neuhaus, F. (2016). Image schemas in computational conceptual blending. *Cognitive Systems Research*, 39, 42-57.
- [30] Fillmore, C. J. (1969). Toward a modern theory of case. In D. A. Reidel & S. A. Schane (Eds.), *Modern studies in English* (pp. 361-375). Prentice-Hall.
- [31] Boden, M. A. (2009). Computer models of creativity. *AI Magazine*, 30(3), 23-34.
- [32] Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Proc. 26th international conference on neural information processing Systems* (pp. 3111-3119).