

物語の系統解析に向けた 自動コード化の試み： Semantic Foldingを用いて

中分遙（高知工科大）・佐藤浩輔（明治大）

本発表の概要

本研究の問題：

- 物語の系統分析（分岐分析）をするためのデータ構築の過程で生じるコード化を自動化したい

発表の目的：

- semantic folding(Webber, 2016)とよばれる自然言語処理の手法によって、物語のコード化したので、その結果を発表し、問題点や改善点等を議論したい

内容

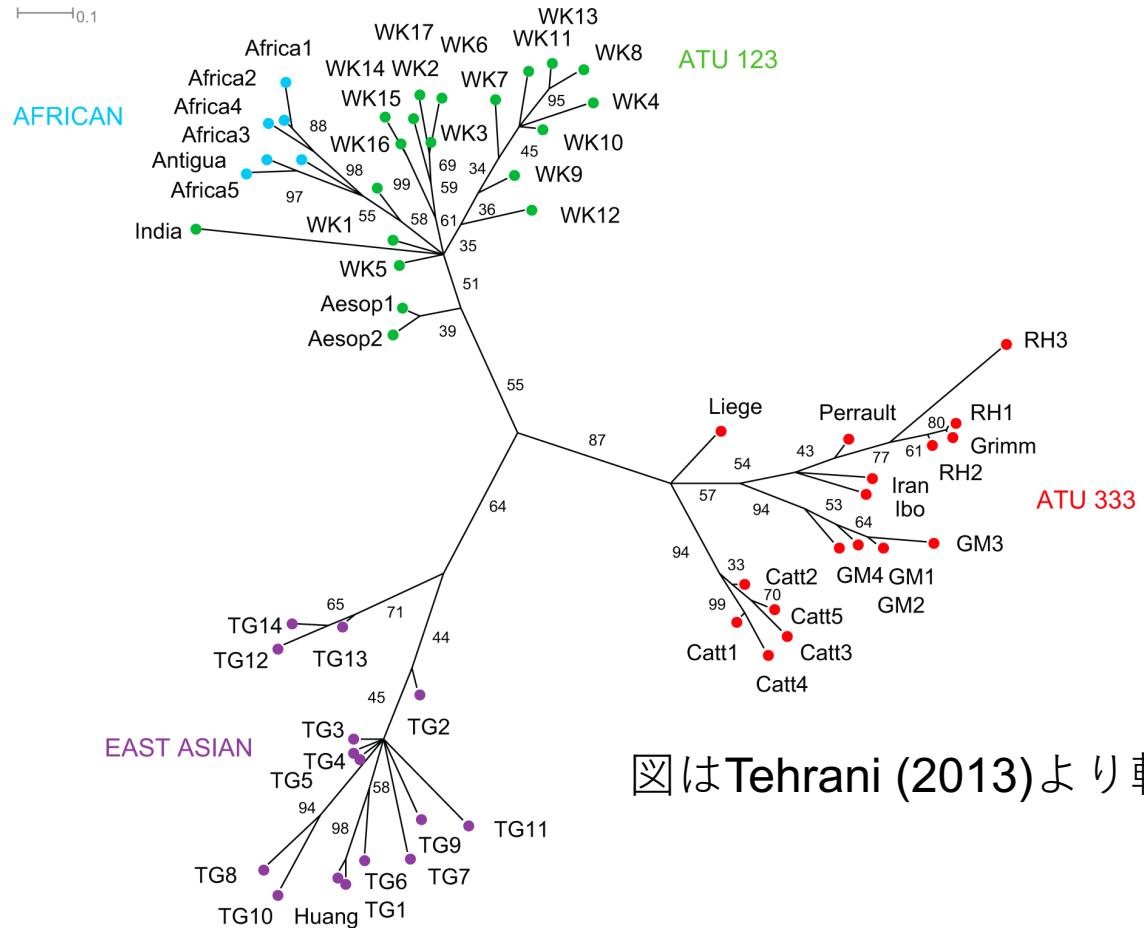
- 複数の物語の内容を手動でコード化したデータで分岐分析を行った結果と自動コード化したデータで分析した結果を比較

物語の文化系統学（文化分岐学）

- 物語のvariationを生物種と見立て、生物系統学の手法を応用することで、物語の系統関係を推測する手法
 - 物語の文化進化研究においては、物語に共通する要素（e.g.,主人公の属性）に基づいて分岐分析を行う研究が多く存在する(for review, Tehrani & d'Huy, 2017¹)
 - 分岐分析：様々な分岐図（系統樹）を作り、最も節約的に説明できる分岐図の同定のため行われる手法（特徴となる形質について、それらが共通の祖先である分岐群から継承したという前提を置く。この形質の変異の数が最小になるように生物種の分岐関係を推定する方法手法）

¹Tehrani, J. J., & d'Huy, J. (2017). Phylogenetics Meets Folklore: Bioinformatics Approaches to the Study of International Folktales. In R. Kenna, M. MacCarron, & P. MacCarron (Eds.), *Maths Meets Myths: Quantitative Approaches to Ancient Narratives* (pp. 91–114). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-39445-9_6

例：「赤ずきんちゃん」に関する研究



図はTehrani (2013)より転載

Tehrani, J. J. (2013). The phylogeny of little red riding hood. *PLoS ONE*, 8(11), e78871.

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0078871>

物語の要素に基づきコード化：
主人公が赤頭巾をかぶっているかどうか
被ってる：1, 被っていない：0

物語1: 1, 0, 1, 1, 0, …, 1

物語2: 0, 0, 0, 1, 1, …, 1

同様に他もコード化
→ このデータ系列に基づき分岐分析

例：都市伝説に関する研究 (Stubbersfield & Tehrani, 2013)

- Stubbersfield & Tehrani (2013)で用いられた都市伝説「ブラディ・メアリー」の類話45話をコード化し分岐分析
 - 物語に含まれる反直観的な概念の数や保持されやすさについて検討

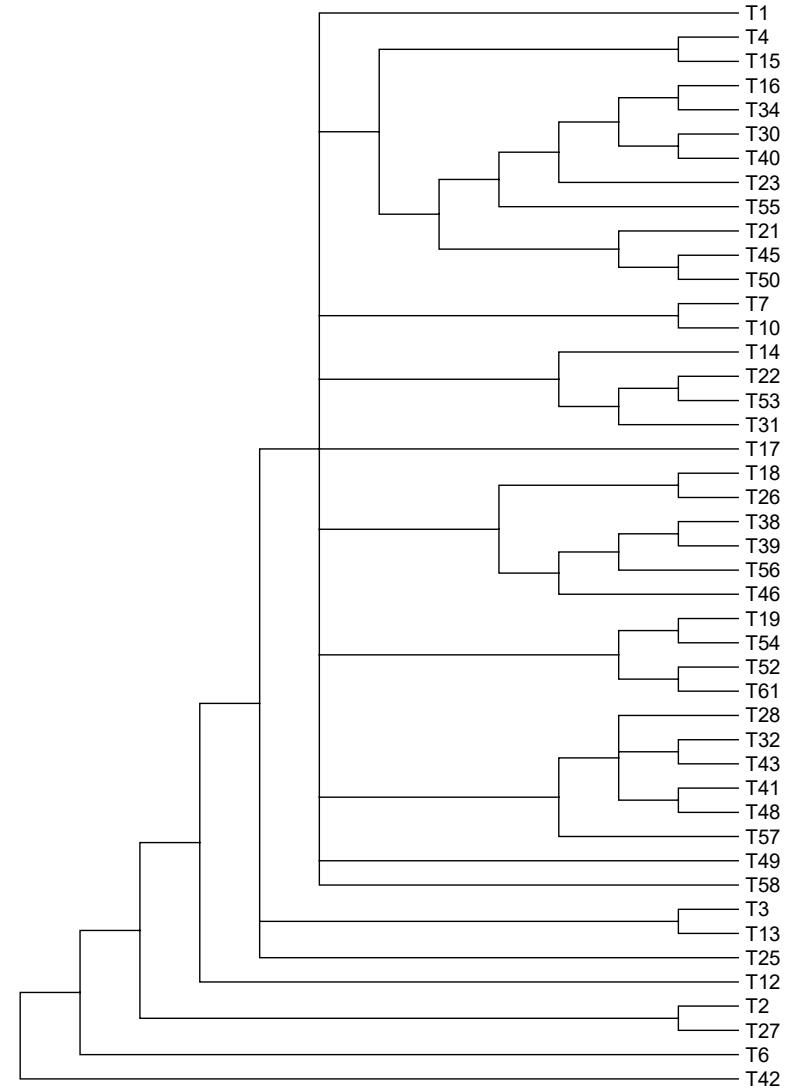
コードの例

Satan or Devilが物語内に出現：有(1) or 無(0)

Bathroomが物語内に出現：有(1) or 無(0)

Stubbersfield, J., & Tehrani, J. (2013). Expect the unexpected? Testing for minimally counterintuitive (MCI) bias in the transmission of contemporary legends: A computational phylogenetic approach. *Social Science Computer Review*, 31(1), 90-102.

<https://journals.sagepub.com/doi/abs/10.1177/0894439312453567>



Stubbers & Tehrani (2013)の図を元に発表者が作成

コード化の課題と自動化の試み

- 紹介した先行研究のような、物語群に対して分岐分析を行うためには、形質（物語の要素）についてコード化する必要がある
- コード化の課題
 - 何が主要な形質なのかを人手で判断しなければならない
 - 客觀性の問題：コード化の基準と判断
 - 民俗学において確立したモチーフ(e.g., Thompson Motif Index)を単位にコード化もできるが、このコード化もトレーニングを要する
 - 大量の物語がある場合に時間が膨大にかかる
 - 民話データはテキスト化されてないものも含めると数が多い
 - 多い場合は一つの話型のコードリストだけで一冊の本になる
- 目的：コード化を自動化することで上記の課題の解決を試みる

コード化の課題と自動化の試み

- 本研究では、Semantic Foldingという手法でコード化
 - Semantic Folding:自然言語文を意味論的特徴に基づいて二値のベクトルとして表現する手法
 - 人文社会学での応用はまだ少数 (e.g., ファイナンス分野では用いられている; Ibriyamova et all 2019)
 - 具体的に人手で意味内容に基づきコード化された場合とどの程度、テキスト間の類似性・距離が一致するのかについては、これまで検討されてこなかった
- そこで本研究では、物語のテキストを手動（人間によって）コード化した Stubbersfield & Tehrani (2013)に着目する
 - このテキストを自動コード化し分岐図を作成：オリジナルと比較

Ibriyamova, F., Kogan, S., Salganik-Shoshan, G., & Stolin, D. (2019). Predicting stock return correlations with brief company descriptions. *Applied Economics*, 51(1), 88-102.

<https://www.tandfonline.com/doi/abs/10.1080/00036846.2018.1494377?journalCode=raec20>

分析

Semantic Foldingを用いた自動的なコード化の試み

- Semantic Folding (Webber, 2016, arXiv):
 - 自然言語のテキストを意味内容を二値ベクトルとして表現する手法。これによって出力された結果は、 semantic fingerprintと呼ばれる
 - 出力された二値ベクトルを「コード化された」テキストとみなす
 - 二値ベクトルはマトリックス形式としても表現できる
 - cortical.ioのAPIを利用してことで、テキストをコード化できる
- 詳細は次ページ以降で
 - cortical.ioによる非常にわかりやすい動画（詳細に関心がある方はyou tubeのリンクを参照ください）
(<https://www.youtube.com/watch?v=HLuRQKzYbb8>)

Webber, F. D. S. (2016). Semantic Folding Theory And its Application in Semantic Fingerprinting.
arXiv. arXiv:1511.08855v2 <https://arxiv.org/abs/1511.08855>
(v1は2015年, v2は2016年)

Semantic Folding (Webber, 2016, arXiv)

1. 意味的マップの設定 (*ソースコードは示されていない)

- スニペットの作成：学習するコーパス（Wikipedia・百科事典等）を意味単位で区切ったもの
- 意味的マップの設定：意味的に似たスニペット群が空間的に近くなるように 128×128 の2次元空間上に配置
- 一つのセルに複数のスニペットが割り当てられている

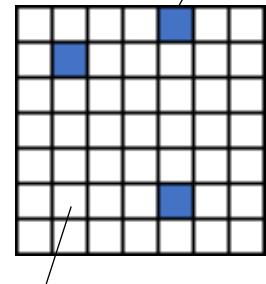
2. 単語のfingerprintの作成

- 例, dogの場合：意味的マップ上でdogを持つスニペットが含まれるセルを1とし、それ以外は0とする。これをdogのfingerprintとする
- 同じようにスニペットに含まれる全ての単語について、fingerprintを作成しデータベースを構築する
 - これら単語のfingerprintsを用いて文のfingerprintも計算できる

dogのfingerprintの
イメージ

Wikipedia dogの記事から
切り取ったスニペット

Domestic **dogs** have
been selectively bred for
millennia ...



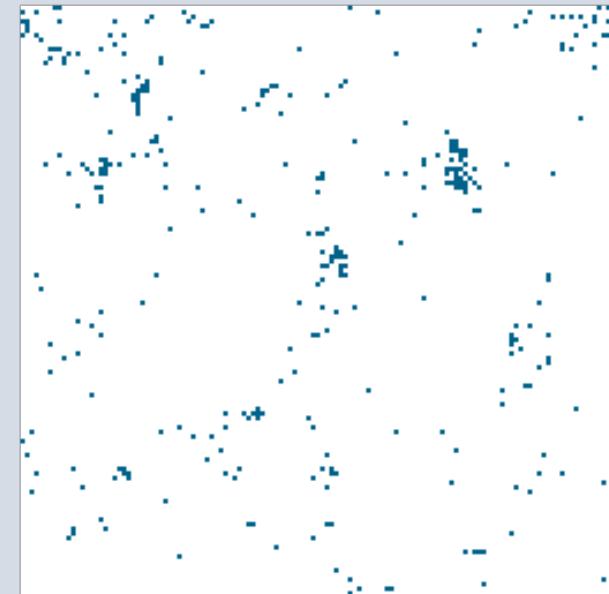
A clade is by definition monophyletic, meaning that it contains one ancestor...

Wikipedia cladeの記事から
切り取ったスニippet

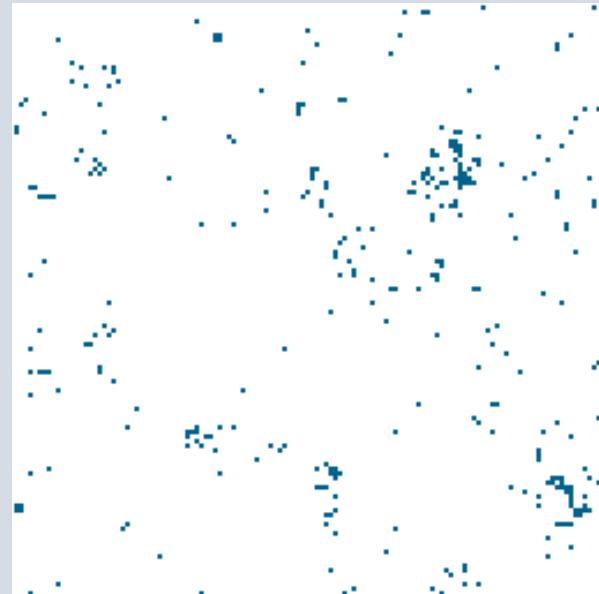
単語のFingerprintの例

: "similar meanings look similar"

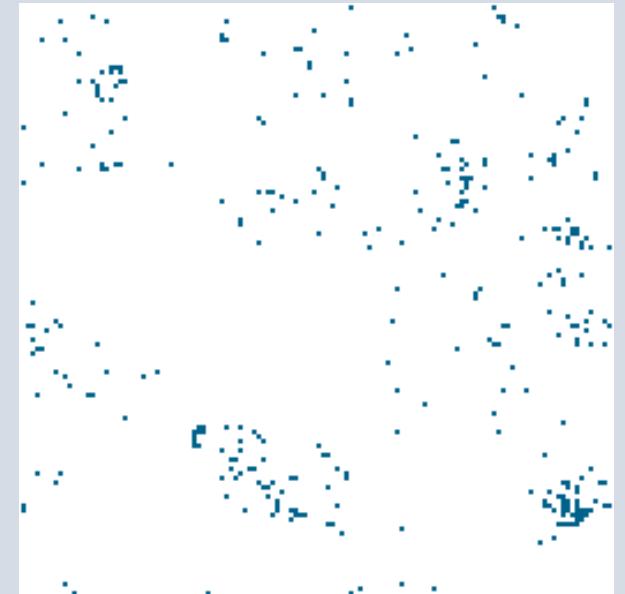
desk



demon



devil



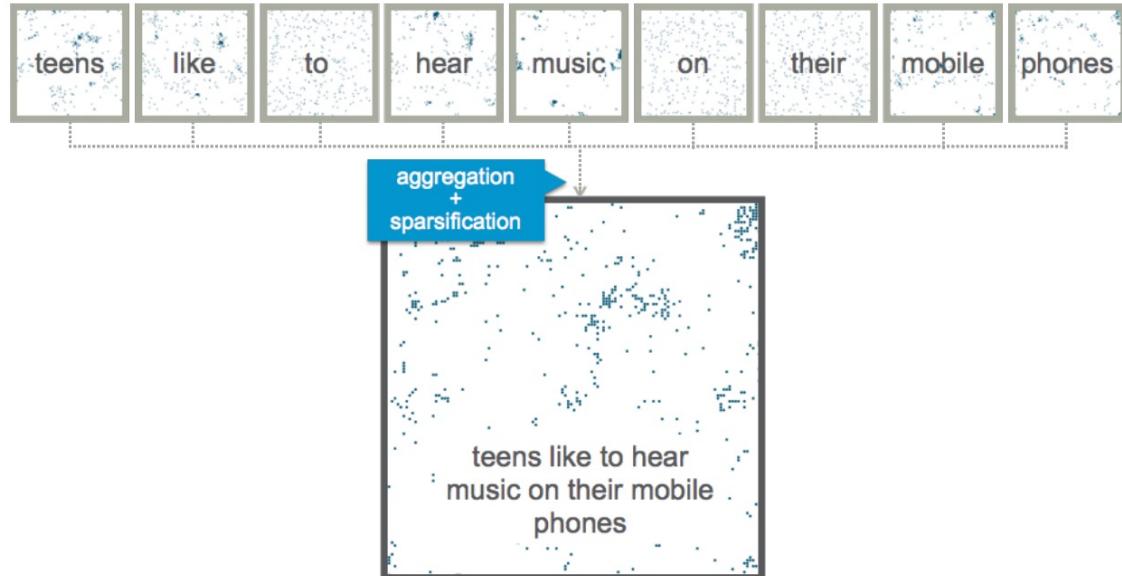
コサイン類似度: 0.19
Euclid距離: 0.81
Jaccard距離: 0.90

コサイン類似度: 0.34
Euclid距離: 0.66
Jaccard距離: 0.78

3. Fingerprintの合成

Fingerprintの合成：

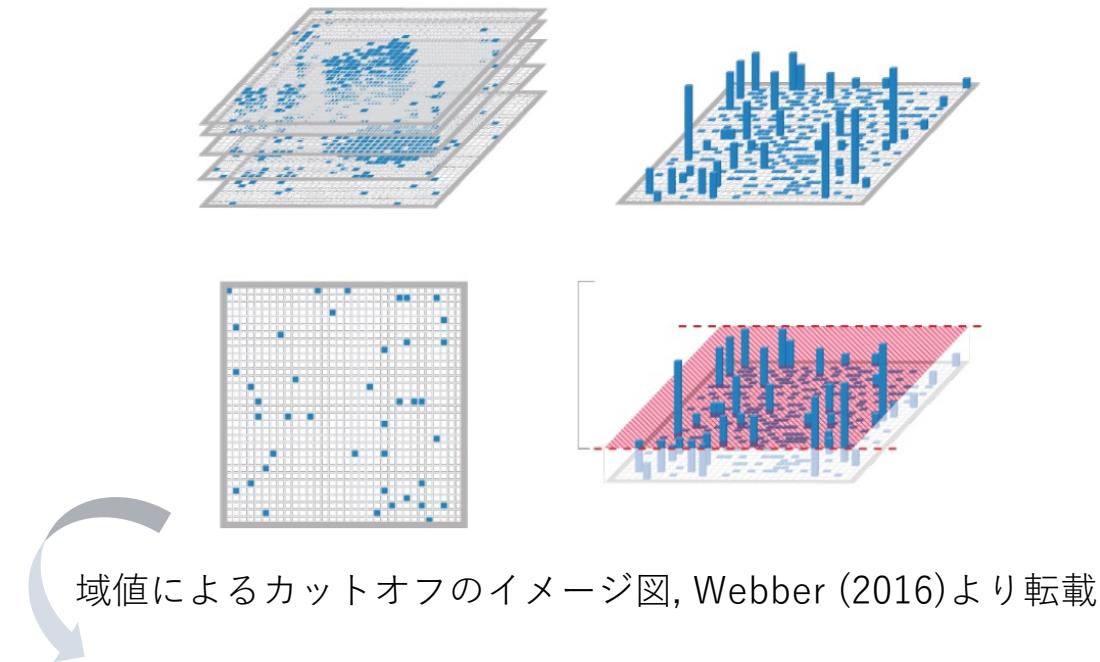
文に含まれる単語のsemantic fingerprintを
全て足し合わせる



Fingerprintの合成のイメージ図, Webber (2016)より転載

本研究では16384個の要素を持つ二値ベクトル
として出力される
(128×128 のマトリックスとしても表現可)

域値によるカットオフ(sparsify)：
足しあわせたfinger printのうち、
上位 984セル → 1
それ以外はのセル → 0



域値によるカットオフのイメージ図, Webber (2016)より転載
この最終的に出力されたものをコード化したものとする
(APIではベクトルとして出力される、これを整列させることでマトリックスとして表現できる)

分析に用いるデータ

- Stubbersfield & Tehrani (2013)で用いられた都市伝説「ブラディ・メアリー」の類話、45話
 - 著者らからテキストデータ・コードデータをシェアしていただきました。
- コードデータ (Semantic foldingには使用しない)
 - 現著者がテキストデータをコード化したもの
 - 36要素（大半が二値だが、二値でないものも含まれる）
- テキストデータ (Semantic foldingに使用)
 - 現著者らがウェブサイトから入手したもの
 - **フォーマットが統一されてない**：例えば物語調 (Once upon a time..)・伝聞調(I was told...)・解説的(This is urban legend...)など

分析で用いるテキストの一例

A long time ago there was a woman named Mary. She was very beautiful and very vain. She had a horrible accident though and her face became so horribly scratched that she bled to death. Her spirit didn't die, she still comes; if you stand in front of a mirror in the dark and say her name three times her face appears. If you do'nt turn on the light and run away as fast as you can she'll scratch your face off, or some say, drag you into the mirror with her!

(do'ntも含め原文のまま)

<http://www.oocities.org/tragicpixie/FLbloodymary.htm>

具体的なテキストのコード化過程

元のテキスト

[A long time ago there was a woman named Mary. She was very beautiful and very vain. She had a horrible accident though and her face became so horribly scratched that she bled to death. Her spirit didn't die, she still comes; if you stand in front of a mirror in the dark and say her name three times her face appears. If you don't turn on the light and run away as fast as you can she'll scratch your face off, or some say, drag you into the mirror with her!]

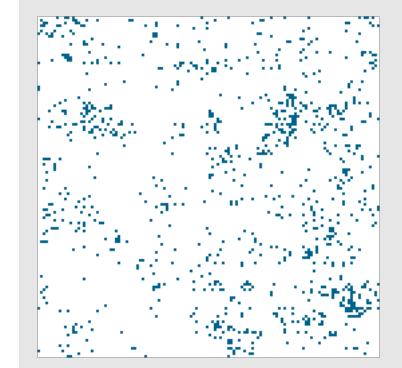
APIによって出力された結果 (984要素)

[7,10,11,13,61,71,78,87,92,101,⋯,16250,
16267,16268,16327]

各要素が含まれるかどうかを0/1に変換
(16384要素)

[0,0,0,0,0,0,1,0,0,1,1,0,1,0,0,0..., 0,0,0]

全ての物語に同様の処理をし、
得られた系列データを分岐分析で用いた
(分岐分析の方法は分析 1 と同一)



系列データを 128×128
で表現した画像

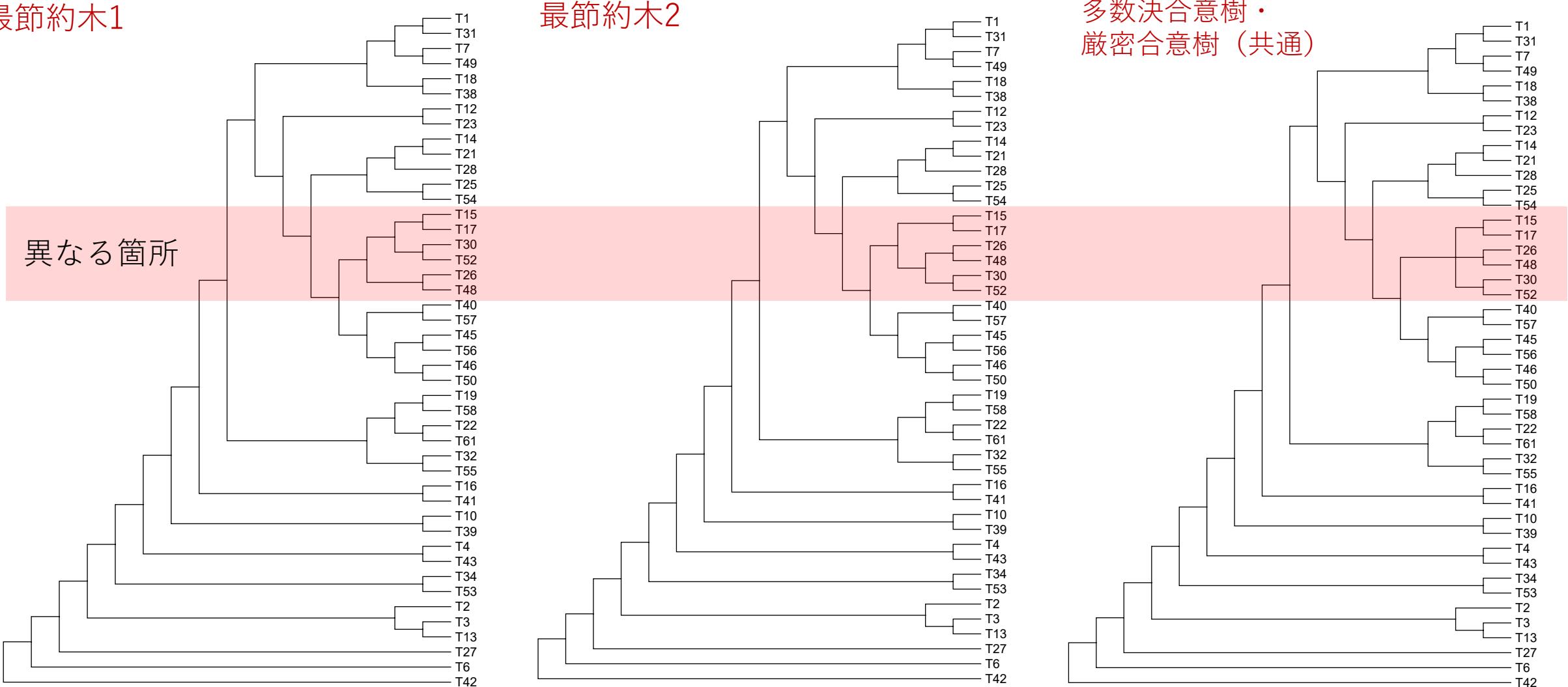
分岐分析

- Stubbersfield & Tehrani (2013) の方法を踏襲
 - 物語数：45
 - 形質数：16384
- 分析ソフト：PAUP4.0
- Heuristic Branch Swapping Searchを用いる
- 外群はT42（伝統的な民話と内容が類似するもの）と設定
- 最節約樹から合意樹（多数決合意樹・厳密合意樹）を作成
 - 先行研究の分析の再現についてはAppendix 1を参照

分岐分析の結果

- 最節約木は2つ：
 - 2つから計算される多数決合意樹・厳密合意樹は一致（2つなためか？）

最節約木1



手動と自動コード化の比較

各話間の距離行列(Jaccard係数)

オリジナル

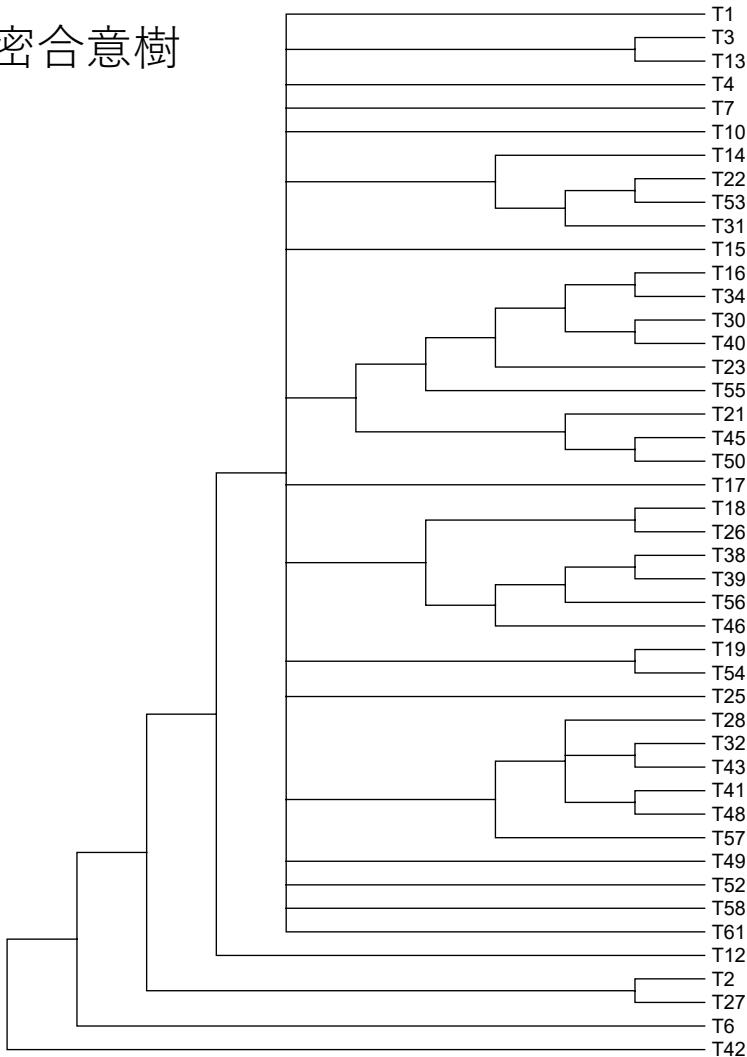
外群であるT42に関しては他の話との
距離が大きい (T6を除く, T6も同じく他との距離が大きい)
T30は他との距離が大きいが分岐図には特徴はない

	T1	T2	T3	T4	T6	T7	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T21	T22	T23	T25	T26	T27	T28	T30	T31	T32	T34	T38	T39	T40	T41	T42	T43	T45	T46	T48	T49	T50	T52	T53	T54	T55	T56	T57	T58	T61
T1	■	0.8	0.6	0.5	0.3	0.6	0.8	0.5	0.4	0.7	0.6	0.7	0.6	0.8	0.6	0.6	0.6	0.8	0.6	0.8	0.6	0.7	0.7	0.8	0.6	1	0.6	0.6	0.5	0.6	0.5	0.8	0.6	0.7	0.4	0.3	0.5	0.5	0.4	0.7						
T2	0.8	■	0.6	0.7	0.8	0.6	0.5	0.6	0.7	0.6	0.7	0.6	0.7	0.7	0.8	0.7	0.8	0.5	0.8	0.8	0.8	0.7	0.8	0.9	0.8	0.8	0.6	0.8	0.6	0.8	0.7	0.8	0.8	0.7	0.7	0.7	0.6	0.7	0.6	0.7						
T3	0.6	0.7	■	0.8	0.7	0.7	0.6	0.6	0.8	0.4	0.5	0.5	0.7	0.7	0.6	0.7	0.7	0.8	0.5	0.7	0.8	0.8	0.7	0.9	0.8	0.9	0.8	0.7	0.8	0.6	0.9	0.6	0.7	0.4	0.7	0.7	0.6	0.7	0.6	0.7						
T4	0.5	0.7	0.1	■	0.5	0.6	0.7	0.6	0.7	0.6	0.8	0.7	0.5	0.8	0.6	0.8	0.9	0.7	0.8	0.7	0.8	0.8	0.9	0.8	1	0.8	0.7	0.7	0.7	0.8	0.6	0.8	0.6	0.8	0.7	0.5	0.6	0.7								
T6	0.8	0.8	0.8	0.9	■	0.9	0.8	0.7	0.7	0.8	0.8	0.7	0.8	0.7	0.9	0.7	0.8	0.9	0.8	0.8	0.9	0.8	0.9	0.9	0.5	0.9	0.9	0.9	0.8	0.9	0.8	0.9	0.8	0.7	0.8	0.7	0.8									
T7	0.3	0.8	0.7	0.6	0.5	■	0.4	0.8	0.7	0.4	0.7	0.6	0.7	0.8	0.6	0.8	0.9	0.7	0.6	0.7	0.6	0.5	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.5	0.4	0.7	0.6	0.7	0.6	0.7	0.7	0.7							
T10	0.6	0.7	0.7	0.5	0.4	0.8	■	0.7	0.6	0.5	0.6	0.6	0.7	0.7	0.6	0.8	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.7	0.5	0.6	0.5	0.7	0.6	0.5	0.4	0.6	0.5	0.6	0.7	0.6	0.7									
T12	0.8	0.8	0.6	0.8	0.7	0.8	0.7	■	0.7	0.4	0.7	0.8	0.4	0.5	0.7	0.9	0.7	0.8	0.6	0.8	0.8	0.9	0.8	0.9	0.8	0.8	0.8	0.8	0.7	0.7	0.6	0.7	0.6	0.7	0.7	0.6	0.7									
T13	0.5	0.6	0.8	0.8	0.7	0.6	0.7	0.6	0.7	0.8	0.6	0.6	0.8	0.6	0.7	0.6	0.7	0.8	0.7	0.8	0.7	0.8	0.6	0.9	0.7	0.5	0.7	0.7	0.6	0.6	0.7	0.6	0.7	0.6	0.7											
T14	0.7	0.6	0.6	0.7	0.4	0.4	0.6	0.5	0.6	0.5	0.6	0.7	0.4	0.5	0.5	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.4	0.8	0.5	0.5	0.4	0.6	0.7	0.4	0.5	0.6	0.7	0.6	0.7										
T15	0.7	0.7	0.6	0.8	0.7	0.6	0.7	0.6	0.5	0.6	0.6	0.5	0.7	0.8	0.7	0.8	0.9	0.8	0.7	0.8	0.7	0.8	0.9	0.8	0.5	0.7	0.5	0.7	0.6	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5									
T16	0.7	0.6	0.8	0.7	0.8	0.7	0.8	0.6	0.6	0.7	0.7	0.6	0.7	0.6	0.7	0.7	0.8	0.7	0.7	0.6	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7										
T17	0.6	0.7	0.4	0.6	0.7	0.6	0.5	0.6	0.7	0.5	0.7	0.6	0.5	0.6	0.7	0.8	0.7	0.9	0.6	0.7	0.8	0.6	0.5	0.6	0.3	0.8	0.4	0.3	0.5	0.4	0.5	0.3	0.4	0.5	0.4	0.5										
T18	0.7	0.6	0.5	0.8	0.7	0.6	0.5	0.6	0.7	0.6	0.7	0.8	0.7	0.6	0.7	0.8	0.9	0.8	0.7	0.6	0.7	0.4	0.6	0.5	0.8	0.7	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5											
T19	0.6	0.8	0.5	0.6	0.7	0.7	0.7	0.6	0.5	0.7	0.4	0.5	0.6	0.7	0.5	0.6	0.7	0.8	0.6	0.7	0.6	0.5	0.6	0.2	0.6	0.7	0.5	0.6	0.5	0.4	0.5	0.6	0.5	0.6	0.5											
T21	0.8	0.8	0.8	0.9	0.8	0.9	0.8	0.7	0.8	0.7	0.7	0.8	0.9	0.8	0.9	0.8	0.9	0.7	0.8	0.7	0.8	0.7	0.8	0.9	0.7	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7										
T22	0.6	0.8	0.7	0.8	0.6	0.5	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.4	0.5	0.6	0.5	0.4	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5										
T23	0.6	0.7	0.6	0.7	0.6	0.6	0.5	0.6	0.4	0.7	0.7	0.6	0.7	0.6	0.7	0.5	0.6	0.7	0.8	0.7	0.7	0.6	0.5	0.6	0.4	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5										
T25	0.6	0.7	0.6	0.8	0.7	0.6	0.5	0.6	0.5	0.6	0.7	0.8	0.7	0.6	0.7	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5										
T26	0.6	0.7	0.7	0.8	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7										
T27	0.6	0.5	0.7	0.6	0.8	0.7	0.6	0.7	0.5	0.6	0.8	0.6	0.5	0.7	0.6	0.5	0.7	0.6	0.8	0.7	0.6	0.7	0.7	0.8	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7									
T28	0.6	0.8	0.7	0.8	0.6	0.6	0.8	0.7	0.5	0.6	0.7	0.4	0.5	0.6	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.4	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6										
T30	0.9	0.8	0.9	0.9	0.9	0.8	0.8	0.9	0.8	0.7	0.7	0.5	0.6	0.7	0.8	0.7	0.8	0.7	0.8	0.9	0.7	0.8	0.8	0.9	0.8	0.7	0.8	0.9	0.8	0.7	0.8	0.9	0.8	0.7	0.8	0.9										
T31	0.8	0.5	0.7	0.8	0.7	0.6	0.5	0.6	0.5	0.6	0.6	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6									
T32	0.6	0.8	0.7	0.8	0.6	0.7	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6									
T34	0.8	0.7	0.8	0.7	0.8	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.6	0.7	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.6									
T38	0.7	0.8	0.8	0.7	0.8	0.7	0.6	0.7	0.5	0.6	0.8	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6									
T39	0.7	0.8	0.7	0.8	0.7	0.8	0.6	0.7	0.5	0.6	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.6	0.7	0.8	0.7	0.6	0.7									
T40	0.8	0.7	0.9	0.9	0.9	0.8	0.7	0.8	0.6	0.8	0.5	0.6	0.7	0.5	0.6	0.7	0.6	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.9	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8										
T41	0.6	0.9	0.8	0.5	0.6	0.7	0.8	0.7	0.9	0.7	0.6	0.5	0.6	0.7	0.8	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.7	0.8	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.6									
T42	1	0.9	0.9	1	0.5	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9								
T43	0.6	0.8	0.7	0.8	0.6	0.7	0.6	0.7	0.5	0.6	0.7	0.4	0.5	0.6	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6									
T45	0.6	0.8	0.7	0.5	0.6	0.8	0.6	0.7	0.8	0.7	0.6	0.7	0.5	0.6	0.8	0.6	0.9	0.5	0.6	0.5	0.6	0.4	0.5	0.6	0.7	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.7	0.6	0.5											

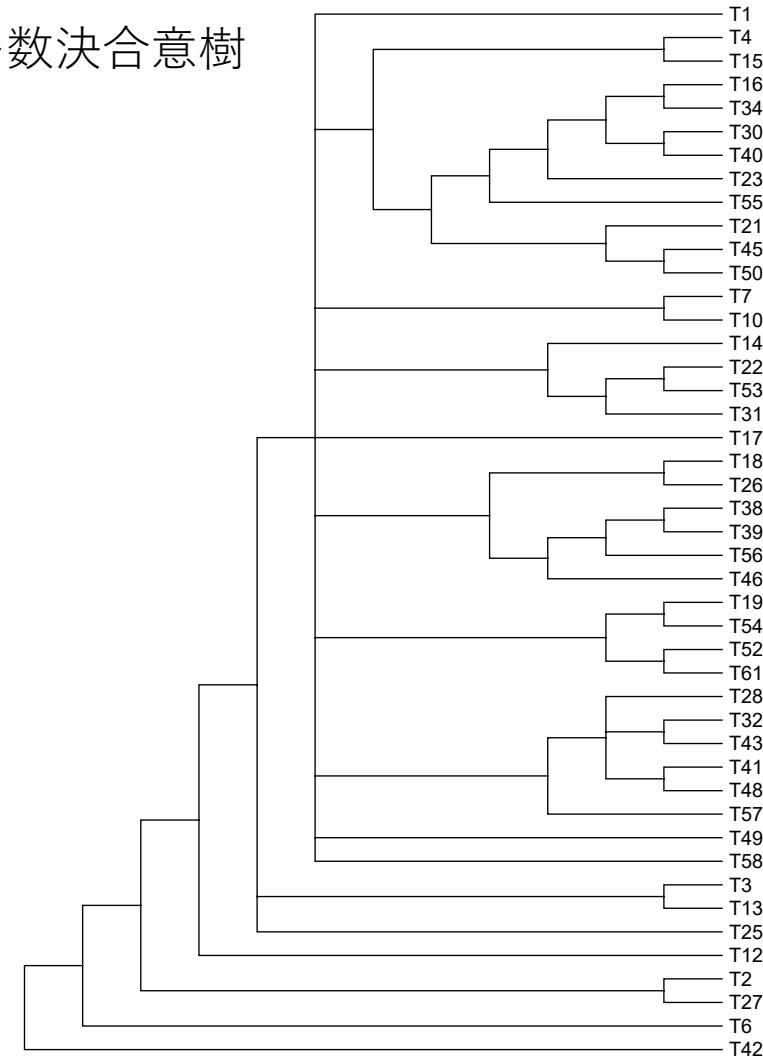
分岐図の比較

オリジナルの系統樹

厳密合意樹

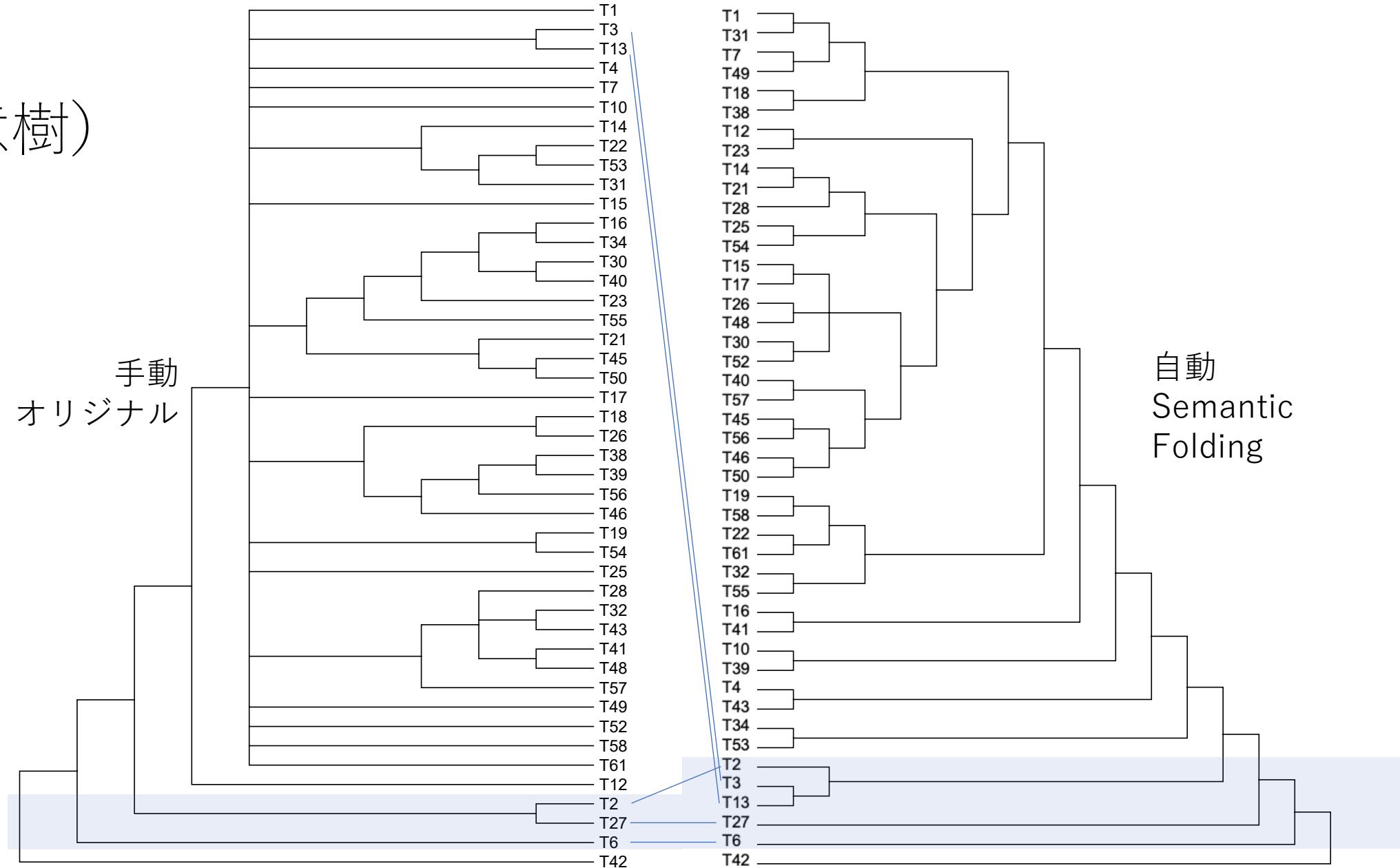


多数決合意樹



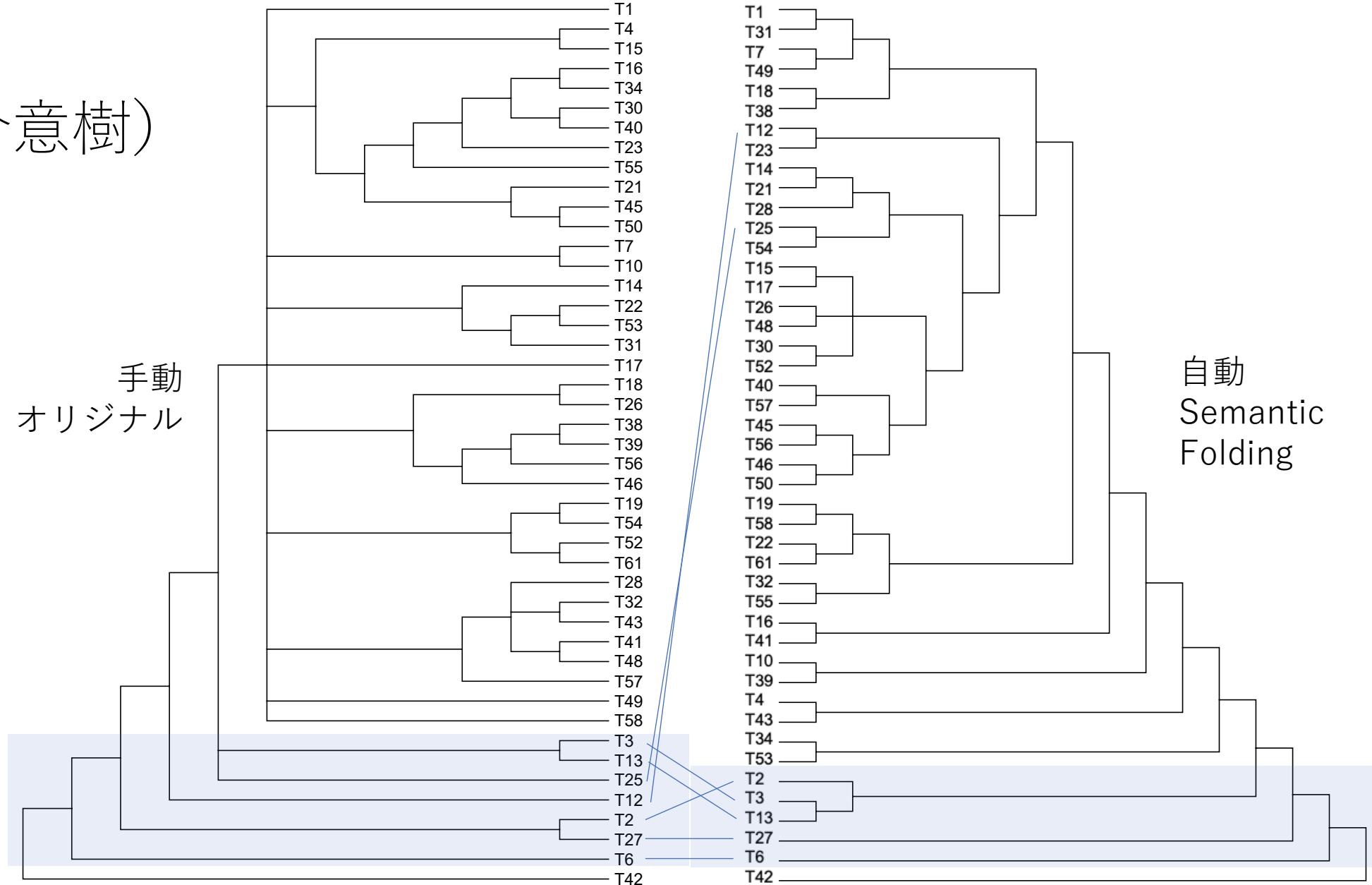
厳密合意樹と多数決合意樹を本研究の結果と比較する

比較 (厳密合意樹)



あまりうまくいっているようには見えないが、外群から早い段階で分岐したものについては一部類似

比較 (多数決合意樹)



あまりうまくいっているようには見えないが、外群から早い段階で分岐したものについては一部類似

結果

- 全体的に再現できることは言い難い
 - ただし、オリジナルの二つの合意樹にも差異はみられる
 - 二つの合意樹で比較的に一貫している初期に外群から分岐したケースについては、本研究ともある程度は一致
- 外群(T42)から、早期に分岐したものについては一致
 - 特に、T6が最初に分岐している点は興味深い
 - T6とT42には、どのような共通点があり、semantic foldingは正しくこの共通点を意味によって識別できているのであろうか？

T42

this is an urban legend which is common in our area. the title says it all you need to stand in a dark room in front of a mirror with a candle held in your hand. you then look into the mirror and say the lords prayer backwards when you have said the lords prayer the devil will be in the mirror behind you.

- devilが含まれる
- Lords prayerといったキリスト教関連の単語が含まれる

T6

I was told that if you said “Hell Mary” seven times in front of a mirror in a dark room, you should see Satan’s image in the mirror. The story was embellished further by the teller, who claimed that after three “Hell Mary”, the mirror turned red, and that after five an unclear face appeared.

- Satanが含まれる
- “Hell Mary”はT6のみ BloodyとHell

・それ以外にも、他の物語には登場するが、両者には登場しないものが多数存在する(e.g., bathroom)。

T42とT6の類似点

- 類似点を探索的に検討
 - オリジンナのコードデータを二値に変換し以下の計算を行う(0 or not 0[1])
 - $E(T42 + T6) - E$ (それ以外の最節約木) E は () 内の平均を計算することを意味
 - この数値の絶対値が大きいものをT42とT6の共通の特徴とする
 - 結果
 - T42とT6のみの特徴として"devil"ないしは"Satan"が出現するという特徴が見られた
 - T42ではdevilが一回用いられる
 - T6ではSatanが一回用いられる
 - また、T42, T6, T2, T12にのみ共通する特徴として、Bloody Mary/Mary Worthが出現しないという特徴が見られた

オリジナルと結果が異なるケース

T12 オリジナルの分析では、早い段階で分岐していたが、本分析では再現されなかった。

A girl named veronica went mad after being molested and horribly abused by her father for years...she went mad from the process and killed herself with a pair of scissors. the legend is if you for Veronica three times while in front of a mirror at midnight, you will see her there, with the scissors still sticking into her neck

コード化されていないが、主人公がVeronicaだったり、家庭内トラブル等かなり特徴的な内容を持つ

Bloody Maryが含まれないという点が外群と類似する
また、この話も文章が短く、他の話型にみられる要素は
含まれていなかった(e.g., bathroom)など

今後の分析課題・可能性

類似・乖離がどのように起きているのか客観的に明らかにする

- テキストをシステムティックに書き換え再分析する
 - 例：devil, Satan, Bloody Mary, bathroomといった物語の特徴であると推定される要素を書き換える・書き加えることで分岐図の変化を見る
- fingerprint自体を計算する
 - 例：「bathroomが含まれる話のfingerprintの合成」から「含まれない話のfingerprintの合成」の差分をとることで、そのfingerprintの形がbathroomの単語と類似するのか検討できる。
 - これによってオリジナルでbathroomが含まれるといったコードされたものから一貫した特徴量が抽出できるのか検討できる
(脳イメージングにおけるtask-specific activation, decodingと類似した発想)

考察：

- 結果：外群との初期の分岐に関しては、一貫する部分はあるものの、全体的な精度はいまいち（ただし、壊滅的なほどでもない）
 - ただし、元の論文のsignalもそこまで強くない点は留意する必要がある
 - またインターネットから収集した資料であり、文の長さやフォーマットにかなり差があったことも留意する必要がある
- 特質すべきは、特に出現頻度が1である、“Satan”や “Devil”に基づき分岐を作ったように解釈できる点
 - これは単語の一致度や出現頻度に基づいた自然言語処理では困難かもしれない、意味を用いたsemantic foldingの利点が出たように見える

考察：

その他Semantic Foldingが持つ特徴の懸念点

- 各物語が持つことができる「形質」の数は固定されている
 - 既存の物語Aをさらに加筆した物語Bがあったとき、物語Aの形質は全て物語Bに継承されているはずだが、形質数が固定されているため、何かしら欠落（押し出される）する要素がある
 - これが、どの程度、分岐分析の結果に影響を与えるのかは明らかでない
- 「形質」数が多く、手動コーディングと比べて分岐が細かくなる
 - 本質的ではない、差異に反応している可能性がある

今後の改善策や方針

- 意味空間 (semantic space) を定義する際に学習するデータを民話を含める
- フォーマットが整った民話集などを利用する
 - 冒頭で紹介したように、赤ずきんちゃんの分岐分析を行った研究も存在するが(Tehrani, 2013)、多言語であるため難しい
 - 一旦、分岐分析から離れて、民話に既にタグづけされているモチーフ、型、カテゴリーが推定できるのか検討すべき？

謝辞

- 本研究のテキストデータ・コードデータを提供してくださった
 - Dr Joseph Stubbersfield (Heriot-Watt University)
 - Dr Jamie Tehrani (Durham University)の2名に感謝いたします。

Appendix

先行研究（手動コーディング）の分岐分析の再現

- （自動化コーディングの前に）まずは、現著らがコード化した系列データを元に分岐分析を行った
 - ただし、PAUP4.0で原著者が分析で用いた設定やパラメータが論文情報に含まれていなかったため、試行錯誤し、原著者らの分析方法を模索
- 原著者らの論文の情報：
 - PAUP4.0を用いた
 - Heuristic Branch Swapping Search（アルゴリズムは不明）
 - T42を外群と設定する
 - 最節約樹からブートストラップ法で合意樹（多数決合意樹・厳密合意樹）を作成
→ 試行錯誤の結果：多数決合意樹・厳密合意樹が再現された（以後同じ設定を採用する）

各話間の距離行列(Dice係数)

オリジナル

外群であるT42に関しては他の話との距離が大きい（T6を除く）
T6も同じく他との距離が大きい

T6 T42

T1	T2	T3	T4	T6	T7	T10	T12	T13	T14	T15	T16	T17	T18	T19	T21	T22	T23	T25	T26	T27	T28	T30	T31	T32	T34	T38	T39	T40	T41	T42	T43	T45	T46	T48	T49	T50	T52	T53	T54	T55	T56	T57	T58	T61
T1	0.7	0.4	0.4	0.8	0.4	0.6	0.3	0.3	0.5	0.5	0.4	0.5	0.4	0.5	0.5	0.5	0.6	0.5	0.6	0.5	0.6	0.8	0.4	0.5	0.5	0.5	0.6	1	0.4	0.4	0.3	0.4	0.5	0.6	0.4	0.5	0.2	0.4	0.3	0.5				
T2	0.7	0.5	0.5	0.6	0.7	0.5	0.4	0.5	0.5	0.5	0.4	0.6	0.7	0.5	0.5	0.5	0.5	0.6	0.5	0.7	0.6	0.6	0.6	0.5	0.7	0.8	0.7	0.8	0.7	0.4	0.5	0.4	0.5	0.7	0.6	0.5	0.6	0.5	0.4	0.5				
T3	0.4	0.5	0.6	0.5	0.5	0.4	0.4	0.5	0.6	0.3	0.4	0.7	0.5	0.6	0.4	0.5	0.7	0.6	0.5	0.8	0.6	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.5	0.9	0.4	0.5	0.5	0.6	0.5	0.5					
T4	0.4	0.5	0.5	0.8	0.4	0.5	0.6	0.7	0.5	0.5	0.5	0.4	0.6	0.5	0.6	0.4	0.6	0.8	0.5	0.6	0.5	0.6	0.7	0.8	0.6	1	0.7	0.5	0.5	0.5	0.5	0.6	0.4	0.4	0.5	0.5	0.5	0.5						
T6	0.8	0.6	0.6	0.2	0.8	0.6	0.5	0.6	0.6	0.5	0.6	0.5	0.8	0.6	0.6	0.8	0.5	0.7	0.7	0.7	0.6	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.1	0.7	0.7	0.6	0.8	0.9	0.5	0.6	0.7				
T7	0.7	0.5	0.4	0.8	0.5	0.6	0.3	0.5	0.5	0.5	0.5	0.4	0.5	0.5	0.5	0.6	0.5	0.8	0.5	0.5	0.5	0.5	0.8	0.5	0.5	1	0.6	0.4	0.3	0.4	0.3	0.4	0.5	0.4	0.5	0.4	0.5	0.3	0.5					
T10	0.4	0.5	0.5	0.6	0.2	0.5	0.5	0.3	0.5	0.4	0.5	0.4	0.4	0.4	0.4	0.5	0.4	0.7	0.5	0.4	0.5	0.4	0.5	0.7	0.6	0.8	0.4	0.4	0.3	0.5	0.3	0.6	0.4	0.5	0.4	0.5	0.5	0.3	0.4					
T12	0.6	0.7	0.4	0.6	0.5	0.6	0.5	0.3	0.7	0.3	0.5	0.8	0.5	0.6	0.5	0.6	0.6	0.8	0.4	0.6	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.7	0.7	0.5	1	0.4	0.5	0.4	0.8	0.5	0.5	0.5						
T13	0.3	0.4	0.7	0.6	0.5	0.5	0.4	0.6	0.4	0.5	0.7	0.5	0.5	0.4	0.5	0.5	0.7	0.4	0.5	0.6	0.6	0.5	0.5	0.7	0.8	0.5	0.6	0.5	0.4	0.5	0.3	0.5	0.6	0.4	0.5	0.4	0.6	0.5	0.4					
T14	0.3	0.4	0.5	0.6	0.3	0.3	0.4	0.4	0.4	0.3	0.4	0.5	0.3	0.3	0.5	0.5	0.4	0.7	0.3	0.4	0.4	0.5	0.7	0.5	0.8	0.5	0.4	0.5	0.2	0.7	0.3	0.3	0.5	0.5	0.2	0.4	0.4	0.4						
T15	0.5	0.5	0.5	0.8	0.6	0.5	0.5	0.6	0.4	0.5	0.4	0.4	0.3	0.4	0.5	0.4	0.5	0.7	0.5	0.5	0.5	0.6	0.7	0.7	0.7	0.8	0.6	0.6	0.5	0.3	0.7	0.4	0.6	0.3	0.5	0.5	0.4	0.4	0.4					
T16	0.5	0.5	0.5	0.6	0.5	0.5	0.7	0.6	0.4	0.5	0.6	0.5	0.5	0.4	0.6	0.4	0.4	0.5	0.4	0.5	0.6	0.5	0.6	0.8	0.5	0.5	0.5	0.4	0.3	0.4	0.4	0.5	0.4	0.5	0.5	0.4	0.5	0.4						
T17	0.5	0.5	0.8	0.6	0.5	0.4	0.3	0.4	0.6	0.2	0.3	0.7	0.4	0.5	0.4	0.7	0.4	0.6	0.5	0.7	0.6	0.5	0.8	0.5	0.5	0.6	0.5	0.4	0.8	0.2	0.5	0.6	0.5	0.2	0.5	0.4	0.3	0.5						
T18	0.5	0.4	0.3	0.5	0.6	0.5	0.3	0.4	0.4	0.5	0.2	0.5	0.5	0.4	0.4	0.5	0.5	0.7	0.5	0.5	0.6	0.5	0.7	0.8	0.6	0.4	0.5	0.3	0.7	0.2	0.4	0.3	0.7	0.5	0.4	0.4	0.4							
T19	0.4	0.5	0.4	0.6	0.5	0.5	0.5	0.4	0.5	0.3	0.5	0.6	0.4	0.4	0.5	0.6	0.7	0.5	0.4	0.5	0.4	0.5	0.7	0.4	0.6	0.8	0.4	0.5	0.6	0.5	0.7	0.5	0.4	0.5	0.3	0.3	0.3							
T21	0.6	0.7	0.5	0.6	0.8	0.6	0.4	0.8	0.7	0.5	0.6	0.7	0.5	0.6	0.8	0.6	0.6	0.7	0.6	0.7	0.6	0.7	0.8	0.7	0.8	0.8	0.6	0.7	0.7	0.5	0.6	0.6	0.8	0.6	0.8	0.5	0.7							
T22	0.4	0.5	0.6	0.5	0.6	0.4	0.5	0.5	0.6	0.5	0.4	0.6	0.5	0.4	0.5	0.6	0.5	0.4	0.3	0.4	0.3	0.4	0.5	0.4	0.5	0.4	0.4	0.3	0.4	0.4	0.5	0.4	0.5	0.3	0.4	0.5	0.3	0.4						
T23	0.5	0.5	0.6	0.5	0.6	0.4	0.5	0.6	0.3	0.4	0.5	0.4	0.5	0.4	0.5	0.6	0.5	0.5	0.6	0.4	0.5	0.6	0.7	0.7	0.8	0.6	0.5	0.5	0.4	0.7	0.5	0.4	0.5	0.4	0.6	0.5	0.4	0.6						
T25	0.5	0.4	0.6	0.6	0.5	0.4	0.5	0.4	0.3	0.4	0.6	0.7	0.5	0.3	0.5	0.3	0.3	0.4	0.4	0.4	0.3	0.4	0.5	0.4	0.4	0.4	0.3	0.5	0.4	0.4	0.4	0.5	0.6	0.5	0.4	0.6	0.5	0.4						
T26	0.5	0.6	0.6	0.8	0.5	0.6	0.5	0.5	0.7	0.6	0.5	0.7	0.6	0.5	0.8	0.7	0.6	0.7	0.5	0.8	0.7	0.7	0.8	0.7	0.8	0.7	0.7	0.7	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6					
T27	0.6	0.3	0.5	0.4	0.5	0.6	0.5	0.5	0.4	0.5	0.6	0.5	0.6	0.4	0.5	0.6	0.5	0.5	0.5	0.4	0.5	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6					
T28	0.5	0.6	0.6	0.7	0.5	0.4	0.5	0.6	0.5	0.4	0.5	0.6	0.5	0.4	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6						
T30	0.8	0.7	0.8	0.8	0.7	0.8	0.7	0.7	0.4	0.7	0.7	0.6	0.5	0.5	0.8	0.5	0.5	0.7	0.3	0.7	0.3	0.7	0.3	0.7	0.8	0.6	0.7	0.7	0.7	0.7	0.6	0.9	0.5	0.7	0.8	0.7	0.8	0.7						
T31	0.4	0.7	0.5	0.8	0.5	0.4	0.3	0.4	0.5	0.5	0.4	0.5	0.4	0.7	0.4	0.6	0.4	0.7	0.5	0.7	0.5	0.4	0.5	0.6	0.5	0.6	0.7	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5					
T32	0.5	0.5	0.6	0.6	0.7	0.5	0.4	0.5	0.6	0.5	0.4	0.5	0.4	0.5	0.6	0.5	0.5	0.6	0.4	0.5	0.5	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.5	0.7	0.5	0.4	0.5	0.4	0.6	0.5	0.4	0.5						
T34	0.5	0.6	0.5	0.6	0.5	0.7	0.6	0.4	0.5	0.6	0.5	0.6	0.4	0.5	0.4	0.6	0.4	0.5	0.4	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5						
T35	0.5	0.6	0.6	0.7	0.8	0.5	0.4	0.3	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.6	0.5	0.7	0.4	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.4	0.5	0.4	0.6	0.5	0.4	0.5					
T36	0.5	0.6	0.6	0.7	0.8	0.5	0.4	0.3	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.6	0.5	0.7	0.4	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.4	0.5	0.4	0.6	0.5	0.4	0.5					
T39	0.5	0.6	0.7	0.8	0.5	0.5	0.7	0.6	0.5	0.6	0.5	0.7	0.5	0.6	0.4	0.5	0.5	0.7	0.5	0.7	0.5	0.6	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8	0.7	0.8			
T40	0.6	0.7	0.6	0.8	0.7	0.8	0.7	0.5	0.7	0.6	0.7	0.4	0.7	0.6	0.8	0.5	0.5	0.6	0.6	0.7	0.7	0.6	0.7	0.7	0.8	0.7	0.8	0.7	0.7	0.7	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.7						
T41	0.5	0.8	0.7	0.6	0.5	0.6	0.7	0.5	0.6	0.5	0.7	0.6	0.5	0.7	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.7	0.5	0.6	0.5	0.6	0.5	0.7	0.5	0.6	0.5	0.6	0.5					
T42	1	0.8	0.8	1	1	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8				
T43	0.4	0.7	0.5	0.7	0.6	0.8	0.4	0.7	0.6	0.4	0.5	0.6	0.5	0.6	0.4	0.5	0.6	0.5	0.7	0.5	0.6	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.7	0.4	0.5	0.4	0.6	0.5	0.4	0.5	0.4	0.6	0.5	0.4	0.5		
T45	0.4	0.7	0.5	0.6	0.8	0.4	0.7	0.6	0.4	0.5	0.6	0.5	0.6	0.4	0.5	0.6	0.5	0.7	0.4	0.5	0.6	0.5	0.6	0.7	0.5	0.6	0.5	0.6	0.7	0.4	0.5	0.4	0.6	0.5	0.4	0.5	0.4	0.6	0.5	0.4	0.5			
T46	0.3	0.6	0.5	0.8	0.3	0.7	0.6	0.5	0.4	0.5	0.6	0.5	0.4	0.5	0.6	0.5	0.4	0.5	0.6	0.5	0.6	0.5	0.6	0.8	0.5	0.6	0.5	0.6	0.5	0.6	0.4	0.5	0.4	0.6	0.5	0.4	0.5	0.4	0.6	0.5	0.4	0.5		
T48	0.4	0.6	0.5	0.6	0.8	0.4	0.5	0.7	0.6	0.5	0.6	0.7	0.4	0.6	0.5	0.6	0.4	0.5	0.6	0.5	0.6	0.5	0.6	0.7	0.6	0.5	0.6	0.5	0.6	0.4	0.5	0.4	0.6	0.5										

Semantic Folding

外群であるT42と他との距離が顕著でなくなる
T12~14, T31の他との距離が顕著となる

T12~14 ↓ T31 ↓ T42

距離
□ 大
■ 小

各話間の距離行列(Jaccard係数)

オリジナル

外群であるT42に関しては他の話との距離が大きい（T6を除く）
T6も同じく他との距離が大きい

T6 T42

Semantic Folding

外群であるT42と他との距離が顕著でなくなる
T12~14, T31の他との距離が顕著となる

T12~14

T31 T42

距離
大
小

各話間の距離行列(Dice係数)

オリジナル

外群であるT42に関しては他の話との
距離が大きい (T6を除く)
T6も同じく他との距離が大きい

T6

T42

	T1	T2	T3	T4	T6	T7	T10	T12	T13	T14	T15	T16	T17	T18	T19	T21	T22	T23	T25	T26	T27	T28	T30	T31	T32	T34	T38	T39	T40	T41	T42	T43	T45	T46	T48	T49	T50	T52	T53	T54	T55	T56	T57	T58	T61
T1	0	0.7	0.4	0.8	0.2	0.4	0.3	0.3	0.5	0.5	0.5	0.4	0.6	0.5	0.5	0.6	0.5	0.5	0.6	0.5	0.6	0.8	0.7	0.7	0.4	0.3	0.4	0.3	0.6	0.4	0.5	0.2	0.2	0.4	0.3	0.3	0.5								
T2	0.7	0	0.5	0.5	0.5	0.7	0.4	0.5	0.6	0.5	0.4	0.6	0.5	0.7	0.6	0.5	0.6	0.6	0.6	0.7	0.8	0.8	0.7	0.7	0.4	0.6	0.5	0.5	0.6	0.7	0.6	0.5	0.5	0.5	0.5	0.5	0.5								
T3	0.4	0.5	0	0.5	0.5	0.5	0.4	0.1	0.4	0.5	0.6	0.3	0.3	0.4	0.7	0.5	0.6	0.4	0.6	0.5	0.7	0.8	0.5	0.6	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.4	0.5	0.4	0.5	0.4	0.5							
T4	0.4	0.5	0.5	0.8	0.4	0.5	0.6	0.7	0.5	0.3	0.5	0.5	0.4	0.6	0.5	0.6	0.8	0.5	0.5	0.7	0.8	1	0.7	0.5	0.6	0.5	0.6	0.4	0.4	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5						
T6	0.8	0.6	0.6	0.8	0.8	0.6	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.8	0.7	0.8	0.8	0.8	0.8	0.6	1	0.7	0.7	0.6	0.8	0.9	0.6	0.6	0.7	0.7						
T7	0.2	0.7	0.5	0.4	0.8	0	0.2	0.6	0.5	0.3	0.5	0.5	0.5	0.5	0.6	0.4	0.5	0.5	0.5	0.5	0.5	1	0.6	0.4	0.3	0.4	0.3	0.6	0.4	0.5	0.4	0.5	0.3	0.5	0.5	0.5	0.5	0.5	0.5	0.5					
T10	0.4	0.5	0.5	0.6	0.2	0	0.5	0.3	0.5	0.4	0.5	0.4	0.4	0.4	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.4	0.4	0.3	0.5	0.4	0.5	0.4	0.5	0.5	0.3	0.4	0.4	0.5	0.4	0.5	0.5	0.5							
T12	0.6	0.7	0.4	0.6	0.5	0.5	0	0.5	0.3	0.7	0.3	0.3	0.5	0.8	0.5	0.6	0.5	0.6	0.5	0.6	0.6	0.7	0.7	0.5	1	0.4	0.5	0.4	0.8	0.8	0.5	0.5	0.5	0.5	0.5	0.5	0.5								
T13	0.3	0.4	0.1	0.7	0.5	0.5	0.5	0.5	0.4	0.6	0.4	0.4	0.5	0.7	0.5	0.6	0.6	0.5	0.8	0.5	0.6	0.3	0.5	0.6	0.4	0.4	0.6	0.5	0.6	0.4	0.5	0.6	0.5	0.5	0.5	0.5	0.5								
T14	0.3	0.5	0.4	0.5	0.6	0.3	0.3	0.4	0	0.4	0.4	0.3	0.4	0.5	0.3	0.5	0.4	0.7	0.3	0.4	0.4	0.5	0.7	0.8	0.5	0.4	0.5	0.2	0.7	0.3	0.3	0.5	0.5	0.2	0.2	0.4	0.4	0.4							
T15	0.5	0.6	0.5	0.3	0.5	0.5	0.5	0.6	0.4	0.5	0.4	0.5	0.4	0.5	0.5	0.7	0.5	0.7	0.5	0.5	0.6	0.7	0.7	0.8	0.6	0.5	0.6	0.3	0.5	0.6	0.4	0.4	0.4	0.4	0.4	0.4	0.4								
T16	0.5	0.6	0.5	0.6	0.5	0.5	0.7	0.6	0.4	0.5	0.6	0.5	0.6	0.4	0.5	0.4	0.5	0.6	0.5	0.5	0.4	0.3	0.4	0.5	0.4	0.6	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5								
T17	0.5	0.3	0.5	0.6	0.5	0.4	0.3	0.4	0.6	0	0.2	0.3	0.7	0.4	0.5	0.3	0.5	0.4	0.7	0.4	0.6	0.5	0.6	0.4	0.2	0.4	0.1	0.6	0.6	0.2	0.3	0.3	0.2	0.3	0.3	0.2	0.3								
T18	0.5	0.4	0.3	0.5	0.6	0.5	0.5	0.3	0.4	0.4	0.5	0.2	0	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.8	0.6	0.6	0.4	0.5	0.3	0.7	0.6	0.4	0.4	0.4	0.4	0.4	0.4	0.4								
T19	0.4	0.6	0.4	0.5	0.5	0.5	0.5	0.4	0.3	0.5	0.5	0.6	0.4	0.5	0.6	0.7	0.4	0.4	0.5	0.6	0.7	0.4	0.5	0.6	0.4	0.6	0.5	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3								
T21	0.6	0.7	0.7	0.6	0.8	0.4	0.8	0.7	0.5	0.4	0.6	0.5	0.7	0.6	0.7	0.5	0.6	0.7	0.6	0.7	0.8	0.6	0.7	0.5	0.5	0.6	0.6	0.5	0.6	0.5	0.7	0.6	0.5	0.7	0.6	0.5	0.7								
T22	0.4	0.6	0.5	0.6	0.4	0.4	0.5	0.5	0.3	0.4	0.5	0.4	0.6	0.3	0.4	0.3	0.5	0.4	0.4	0.8	0.4	0.3	0.5	0.4	0.3	0.4	0.5	0.6	0.3	0.3	0.4	0.3	0.4	0.3	0.4	0.3	0.4	0.3							
T23	0.5	0.5	0.6	0.5	0.4	0.5	0.6	0.4	0.5	0.4	0.5	0.4	0.5	0.5	0.5	0.6	0.5	0.6	0.5	0.6	0.7	0.7	0.8	0.5	0.6	0.5	0.6	0.4	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5								
T25	0.5	0.4	0.6	0.6	0.5	0.4	0.5	0.3	0.4	0.3	0.4	0.3	0.5	0.7	0.6	0.5	0.5	0.6	0.5	0.4	0.4	0.4	0.5	0.3	0.4	0.5	0.4	0.6	0.2	0.3	0.6	0.5	0.4	0.5	0.4	0.5	0.4	0.5							
T26	0.5	0.6	0.6	0.8	0.5	0.6	0.5	0.5	0.7	0.6	0.5	0.4	0.7	0.6	0.7	0.7	0.5	0.6	0.7	1	0.7	0.6	0.4	0.5	0.4	0.7	0.5	0.4	0.6	0.5	0.6	0.4	0.5	0.5	0.6	0.5	0.6	0.5	0.6						
T27	0.6	0.5	0.4	0.5	0.6	0.5	0.5	0.5	0.4	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6								
T28	0.5	0.6	0.7	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.3	0.5	0.5	0.5	0.5	0.2	0.4	0.4	0.5	0.3	0.4	0.5	0.4	0.4	0.2	0.3	0.5	0.6	0.5	0.7	0.2	0.3	0.5	0.6	0.5	0.7	0.3								
T30	0.8	0.7	0.8	0.8	0.7	0.8	0.7	0.7	0.7	0.6	0.5	0.5	0.5	0.7	0.7	0.8	0.7	0.7	0.8	0.6	0.7	0.7	0.7	0.6	0.7	0.6	0.7	0.6	0.7	0.5	0.7	0.8	0.7	0.6	0.7	0.5	0.7								
T31	0.4	0.7	0.3	0.5	0.7	0.5	0.5	0.4	0.3	0.5	0.4	0.5	0.4	0.7	0.4	0.5	0.7	0.5	0.5	0.4	0.5	0.4	0.5	0.5	0.4	0.5	0.5	0.5	0.4	0.4	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5							
T32	0.5	0.6	0.6	0.7	0.5	0.4	0.6	0.4	0.5	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5							
T34	0.5	0.6	0.5	0.6	0.5	0.7	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5							
T38	0.5	0.6	0.6	0.7	0.5	0.4	0.6	0.5	0.5	0.7	0.5	0.4	0.7	0.4	0.5	0.5	0.6	0.5	0.5	0.7	0.4	0.3	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4						
T39	0.5	0.6	0.7	0.5	0.5	0.7	0.6	0.5	0.6	0.5	0.5	0.6	0.5	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5							
T40	0.6	0.7	0.6	0.8	0.7	0.8	0.7	0.7	0.6	0.7	0.4	0.7	0.4	0.7	0.3	0.8	0.5	0.7	0.8	0.7	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6									
T41	0.5	0.8	0.7	0.6	0.8	0.7	0.6	0.7	0.5	0.6	0.7	0.3	0.7	0.6	0.7	0.6	0.7	0.5	0.6	0.7	0.4	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6	0.5	0.6								
T42	1	0.8	0.8	1	0.3	1	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8							
T43	0.4	0.7	0.6	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4	0.5							
T45	0.4	0.7	0.6	0.5	0.4	0.5	0.4</td																																						