

多言語用例対訳グラフにおける メタノード作成手法の提案と評価

福島 拓^{†1} 吉野 孝^{†2}

世界的なグローバル化が進んでおり、医療分野では多言語用例対訳を用いた外国人患者の支援が行われている。そこで、我々は多言語用例対訳共有システムの開発を行っている。共有対象の用例対訳は、各言語間で一対一に対応する必要があるが、用例対訳には各言語間で一対一に対応しない言葉の組み合わせが存在している。このような多言語間の言葉の多様性への対応が求められているが、単純な用例間の意味のつながり情報のみでは対応することができていなかった。そこで、本稿では、一対多、多対多の関係にある用例対訳グラフを提供可能な用例対訳とするために、メタデータ作成手法を提案する。また、手法運用時の問題解決として、新たな用例間リンクの発見手法や、メタノードに意味を持たせるための方法についても提案する。

Proposal and Evaluation of Meta-node Composition Method on Multilingual Parallel-text Graph

TAKU FUKUSHIMA^{†1} and TAKASHI YOSHINO^{†2}

Recently, there is increasingly globalized world. Medical workers use multilingual parallel texts for foreign patients in medical field. We developed a multilingual parallel-text sharing system. A parallel-text requires one-on-one combination among each language. However, parallel texts often have example sentences of many-to-many combination among each language. A multiplicity of expression causes this problem. It is difficult to solve this problem only information of association between example sentences. Therefore, we proposed a meta-node composition method for parallel-text graph having many-to-many combination. Moreover, we proposed a new parallel-text link discovery method and a method of giving meaning to meta-node.

1. はじめに

近年の世界的なグローバル化により多言語間コミュニケーションの機会が増加している。しかし、一般に多言語を十分に習得することは非常に難しく、母語以外の言語によるコミュニケーションは困難なこともあり¹⁾⁻³⁾、日本語を理解できない外国人と日本人とのコミュニケーションは十分に行うことができない。このため、用例対訳や機械翻訳などの言語資源を組み合わせることで利用できる仕組みである言語グリッドの活動が広がるなど^{4),5)}、言語の壁を越える活動が活発化している。

日本語を理解できないことの影響が顕著に現れる分野の1つに医療がある。医療分野では、わずかなコミュニケーション不足で医療ミスが発生する恐れがある。特に、日本語が通じない外国人と日本人の医療従事者間でのやり取りは、意思の疎通を十分に行うことができない。現在、日本語を理解できない外国人の支援は医療通訳者が行っているが、医療通訳者は慢性的な人員不足となっている。また、通訳者の身分保障や通訳者自身のメンタルケアなどの問題が存在している⁶⁾。

情報技術を利用した医療分野の支援として、正確性の確保が可能な用例対訳が多く用いられている。用例対訳とは、用例を多言語に翻訳した多言語コーパスのことを指す。用例対訳は言語間の用例の変換を目的としているため、一つの用例対訳に同じ言語の複数の用例が含まれないという特徴がある。用例対訳を利用したシステムとして、多言語医療受付支援システム M^3 (エムキューブ)⁷⁾ や、多言語問診票作成システム⁸⁾ がある。 M^3 は、用例対訳を用いて医療受付での応対や問診の支援を行っている。また、多言語問診票作成システムは、用例対訳や機械翻訳を用いて医療受付で記入を行う問診票の多言語化を可能としている。

用例対訳は前述の通り、ある言語の用例を他の言語の用例に一意に変換可能であるという特徴がある。しかし、多言語の言葉には各言語間で一対一に対応せず、多対多の組み合わせとなるものが存在している。また、使用する相手やニュアンスの違いによって同じ意味の複数の言葉が存在する場合も多い。このため、用例対訳も一対多、多対多の関係になり得るものが存在している。一対多、多対多の関係にある用例対訳はグラフデータ構造で表すことができるが、用例対訳の特徴である多言語間の言葉の変換を一意に行うことができない。

^{†1} 和歌山大学大学院システム工学研究科

Graduate School of Systems Engineering, Wakayama University

^{†2} 和歌山大学システム工学部

Faculty of Systems Engineering, Wakayama University

このような一意変換できない用例対訳は、多言語対話システムで使用することができない。また、複雑化した用例対訳グラフは、各用例間の関係リンクが不足している場合も発見が難しいという問題がある。これは、医療分野の用例対訳だけでなく、様々な分野の用例対訳に対して言えることである。

そこで本稿では、一対多、多対多の関係にある用例対訳グラフを提供可能な用例対訳とするために、メタデータを利用したデータ構造を提案する。また、用例対訳グラフとメタデータを利用し、不足している各用例間の関係リンク発見手法について述べる。

2. 関連研究

用例対訳の需要増加を背景に、用例対訳の収集が行われている。Chen らは Web 上にある用例対訳を自動的に収集する試みを行っている⁹⁾。また、Bond らは田中コーパス¹⁰⁾を基に、用例対訳の収集プロジェクトを行っている¹¹⁾。このプロジェクトは TATOEB A プロジェクトという名前で活動が行われており、日常的に使用する用例の収集を、日本語、英語、フランス語、中国語、ドイツ語など様々な言語で行っている。我々も正確性が求められる医療分野の用例対訳の収集、共有を目的とした、多言語用例対訳共有システム TackPad の開発を行っている¹²⁾。しかし、これらの研究では多言語間に存在する一対多や多対多の関係について考慮していない。本稿では、多言語用例対訳における一対多や多対多の関係を考慮したメタノード作成手法を提案する。

メタノードの概念を言葉に適用したものとしてシソーラスがある。梶らはコーパスからのシソーラスの自動生成を行っている¹³⁾。また、李は名詞概念と動詞概念をリンクで結合したハイパー・シソーラスを提案している¹⁴⁾。福井らはシソーラスを用いた多言語翻訳を提案している¹⁵⁾。これらは、シソーラスの特徴である上位概念と下位概念を結んだ木構造を基本として研究が行われている。しかし、本稿で扱う多言語用例対訳は用例間の意味のつながり、つまりグラフデータ構造が基本となっている。このため、本稿のメタノード作成手法はグラフデータ構造への適用を行うため、シソーラスとは本質的に異なる。

用例対訳以外では、グラフデータ構造のグループ化が行われている。小島らはハイパーリンクを用いた Web ページ群のグループ化を行っている¹⁶⁾。また、岡田らはソーシャルネットワークや Web への適用を目的としたコミュニティ発見手法を提案している¹⁷⁾。我々もこれらの研究と同様に、グラフデータ構造である用例対訳の分類を行う。その際、メタノードを用いて用例対訳を分類する。さらに、メタノードを用いた用例対訳の提供時における問題についても解決を目指す。

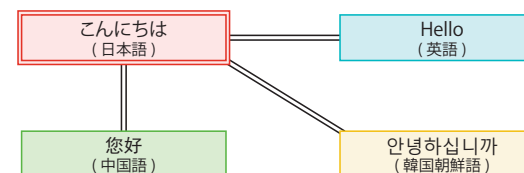


図 1 用例対訳の例
Fig. 1 Example of parallel-text.

3. 多言語用例対訳グラフにおけるメタノード作成手法

本稿で提案するメタノード作成手法は、一対多、多対多の関係性を持つ用例対訳グラフを、多言語対応システムで利用可能にするために用いる。まず、3.1 節で多言語用例対訳における一対多、多対多の問題について述べた後、3.2 節で本手法と、作成したメタノードの提供方法について述べる。さらに、3.3 節で不足している用例間関係リンク発見手法について述べる。

なお、本手法は多言語用例対訳共有システム TackPad¹²⁾ に適用する。

3.1 多言語用例対訳の一対多、多対多における問題点

用例対訳の例を図 1 に示す。用例対訳は、(1) 元となる用例を準備する、(2) 用意された(1) の用例を各言語に翻訳する、という流れで作成することが一般的である。しかし、(2) で作成された用例間で意味が同じであるという担保が行われない。

図 1 を例に説明する。図 1 は元となる用例が日本語で、英語、中国語、韓国朝鮮語の各用例が作成された例である。このとき、意味が同じであると担保されているのは「日本語-英語」「日本語-中国語」「日本語-韓国朝鮮語」の用例間のみであり、「英語-中国語」「英語-韓国朝鮮語」「中国語-韓国朝鮮語」の用例間では意味が同じとは限らない。このため、先ほどの手順の後に、「(2) で作成された翻訳用例間で意味が同じであることを確認する」という手順が必要であると考えられる。このことを明確にするために、我々が開発している TackPad では意味が同じである用例間にリンクの作成を行っている。本稿では、このリンクを「用例間リンク」とする。用例間リンクがつけられた用例は意味が同じであるため、用例対訳での多言語間の言葉の変換で使用することができる。しかし、実際の用例対訳は一対多や多対多の関係を持つものが存在している。多対多の関係性を持つ用例対訳は、一意に言葉の変換を行うことができない。多対多の関係を持つ用例対訳の例を図 2 に示す。図 2 中の二重線は、用例間リンクを示す。なお、図 2-(1) と図 2-(7) の間に用例間リンクを作成することが可能

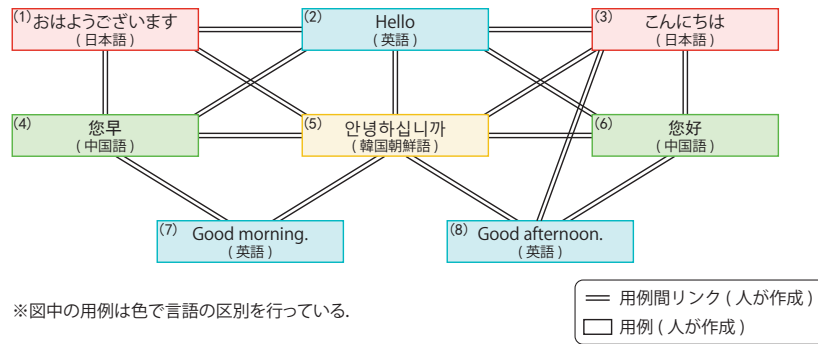


図 2 用例対訳の用例間リンク例
Fig. 2 Example of parallel-text link.

であるが、以降の説明の都合上、用例間リンクがない例を使用している。

図 2-(5) の韓国朝鮮語はどの時間帯でも使用できる挨拶である。このため、図 2 では各言語の朝の挨拶や昼の挨拶である 7 つの用例と用例間リンクでつながっている。図 2 の用例対訳の場合、図 2-(5) の韓国朝鮮語とリンクしている用例が言語ごとに一意でない (英語の 3 用例、日本語と中国語の 2 用例ずつつながっている) ため、簡単には多言語間の言葉の変換を行うことができない。

簡単に多言語変換ができないことは、用例対訳を多言語対応のシステムで使用する際に問題となる。例として多言語対話システムを上げる。多言語対話システムは対話のフローをあらかじめ用意している場合が多い。対話フローの作成には、質問と回答の対を用意する必要がある。質問と回答の対の例として、「この病院は初めてですか?」という質問と「はい」「いいえ」の回答の組み合わせがある。この用例を元に、他の言語へ変換することで多言語対応を行っている。このとき、日本語に対して同じ言語の複数の用例がリンクされている場合、その用例対訳は言葉の一意変換ができない。結果的に、一対多、多対多の関係を持つ用例対訳は多言語対話システムで使用できない。

そこで本稿では、「メタノード」を用意することで、一対多、多対多の関係を持つ用例対訳を一意的に変換可能な一対一の用例対訳へ変換を行う。メタノードの作成については次節で説明を行う。

3.2 メタノードの作成手法

本節では、前節で述べた問題の解決に用いるメタノードの作成手法について説明する。

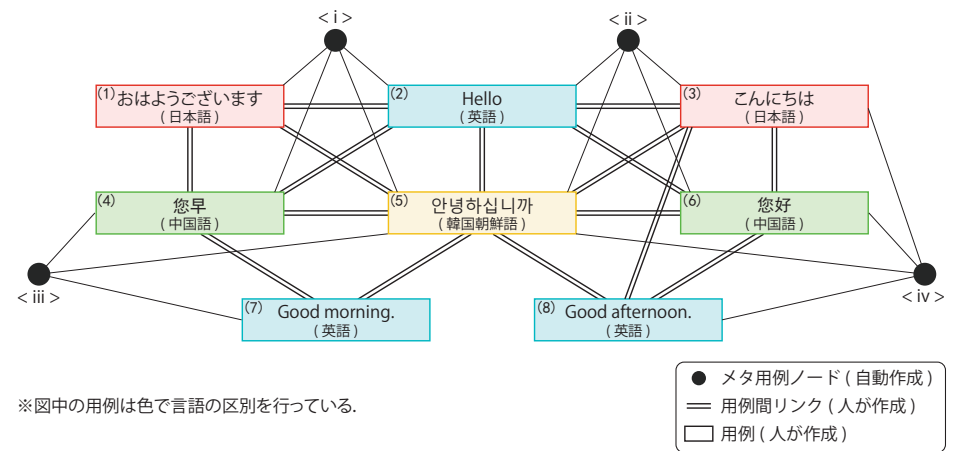


図 3 用例対訳のメタノードの作成例
Fig. 3 Example of meta-node for parallel text.

本稿で作成するメタノードは次の特徴を持つ。

- メタノードは同じ意味の多言語用例を含む
- 1 つのメタノードには同一言語の用例は複数含まない

このことから、本稿では用例対訳グラフの中で完全グラフ (すべてのノードが相互にリンクされているグラフ) となっているものをメタノードとする。なお、メタノードは最大のクリーク (一番大きな完全グラフ) をリンクしたノードとする。メタノードは、利用者が作成する用例間リンクを元に動的に生成するものとする。このことで、用例対訳の作成者には従来から行っている「用例の作成」と「用例間リンクの作成」以外の作業の負担を求めない形で、提供可能な用例対訳を作成する。

図 2 にメタノードを付与した例を図 3 に示す。図 3 の黒丸が作成したメタノードである。メタノード (i),(ii),(iv) は、完全グラフとなっていた 4 つの用例を、メタノード (iii) は、完全グラフとなっていた 3 つの用例をそれぞれ含んでいる。

このようにして作成されたメタノードは、多言語対応システムに用例対訳として提供するとき使用する。例えば、図 3 のメタノード (i) を利用すると、朝の挨拶の用例対訳が、図 3 のメタノード (iv) を利用すると、昼の挨拶の用例対訳がそれぞれ利用できる。

しかし、日本語しか分からない利用者は、メタノード (ii) とメタノード (iv) のどちらの

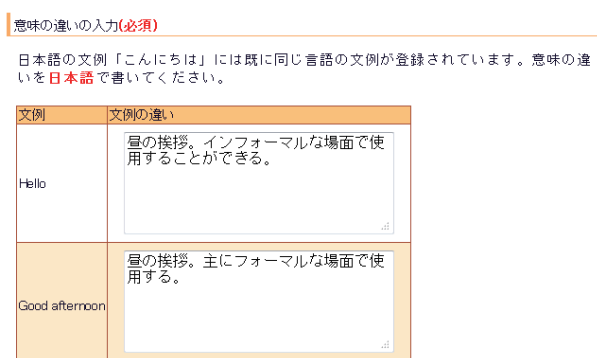


図 4 メタノードに付与する意味の入力画面例
Fig. 4 Screenshot of giving meaning to meta-node function.

「こんにちは」を利用すべきか判断することが難しい。このため、メタノードに意味を持たせることを考える。本手法の適用先である TackPad では、一つの用例に対して、同じ言語の複数の用例が付与されたときに、意味の違いの入力を促している。画面例を図 4 に示す。図 4 は、日本語と英語を理解できる翻訳者が、既に「Hello(図 3-(2))」が付与されている用例「こんにちは(図 3-(3))」に対して、「Good afternoon(図 3-(8))」を付与したときの例である。この状態では、「Hello」と「Good afternoon」の違いが分からない日本語話者はどちらを選択すべきかの判断が難しい。しかし、図 4 のように意味の違いを日本語で記述することで、日本語話者は使用場面に合わせて用例対訳の選択を行うことができる。また、意味の違いをそれぞれのメタノードに付与することで、各言語でメタノードに意味を持たせることができる。このことで、利用者はメタノードの意味を見て、どのメタノード(用例対訳)を使用するか判断することができる。

3.3 不足用例間リンクの発見手法

一対多、多対多の関係にある用例対訳においては、関係が複雑になっているため用例間リンクが不足している場合も発見することが難しい。このため、本節では不足している用例間リンクの発見手法について述べる。本手法で発見した不足用例間リンクは、用例対訳作成システムの利用者に提示する。このようにすることで、新たな用例間リンクの作成補助ができると考えられる。

本手法の手順を以下に示す。

- (1) ある用例 A について着目する。

- (2) 次の条件すべてに合うメタノード α が存在した場合、「用例 A と用例間リンクを介してつながれている用例」と「メタノード α に属している用例」で、共通している用例の数を調べる。
 - (a) メタノード α には用例 A と用例間リンクでつながれている用例のいずれか(もしくは全部)が含まれる。
 - (b) メタノード α には用例 A は属していない。
 - (c) メタノード α に用例 A と同じ言語の用例が属していない。
- (3) 「(メタノード α と用例 A に共通の用例数)/(メタノード α に属する用例数)」を求める。本稿ではこの値を「メタノードとの一致度」とする。メタノードとの一致度が大きいほど用例 A はメタノード α に属する可能性が高くなる。

上記の内容を図 3 を用いて説明する。図 3-(1) は用例間リンクにより用例 (2),(4),(5) とつながっている。用例 (2),(4),(5) が属している用例 (1) が属していないメタノードは、(ii),(iii),(iv) である。しかし、メタノード (ii),(iv) には、用例 (1) と同じ言語である日本語の用例が既に属しているため、用例 (1) は属することができない。このため、メタノード (iii) について確認する。「用例 (1) と用例間リンクを介してつながれている用例」と、「メタノード (iii) に属している用例」とで共通している用例は用例 (4) と (5) である。これらのことから、用例 (1) のメタノード (iii) との一致度は $(2/3)=67\%$ であると求められる。メタノードとの一致度が高いものが不足している用例間リンクである可能性が高いため、効率的に利用者に不足用例間リンクの確認を依頼することができる。今回の例の場合、用例 (1) と用例 (7) に用例間リンクの作成される可能性が利用者に提示されることとなる。

4. 多言語用例対訳共有システムへの適用

多言語用例対訳データ構造を、次に示す用例対訳に適用した。

- (1) 多言語用例対訳共有システム TackPad¹²⁾ で収集済みの用例対訳
- (2) 多言語問診票作成システム⁸⁾ で使用している用例対訳

TackPad の収集言語は日本語、英語、中国語、韓国朝鮮語、ポルトガル語、スペイン語、ベトナム語、タイ語、インドネシア語の 9 言語である。また、収集済みの用例数は 14,487 件、用例間リンク数は 18,285 件であった。多言語問診票作成システムで使用している用例対訳は、日本語、英語、中国語、韓国朝鮮語、ポルトガル語、ベトナム語の 6 言語である。また、用例数は 2,480 件、用例間リンク数は 5,278 件であった。用例対訳ごとの言語別の用例数を表 1 に示す。TackPad の用例対訳は、日本語が他の言語の 2 倍以上収集されている

表 1 各言語の用例数

Table 1 Number of parallel texts in each language.

言語	用例数	
	TackPad	問診票
日本語	4,567	508
英語	2,427	473
中国語	2,330	484
韓国朝鮮語	2,136	436
ポルトガル語	2,248	475
スペイン語	706	0
ベトナム語	26	104
タイ語	38	0
インドネシア語	9	0
合計	14,487	2,480

・表中の「問診票」は多言語問診票作成システムを指す。

表 2 メタノードに結合された用例数

Table 2 Number of example sentences in meta-node.

結合された用例数	メタノード数	
	TackPad	問診票
2	4,252	0
3	37	2
4	5	73
5	1,399	336
6	0	108
合計	5,693	519

・表中の「問診票」は多言語問診票作成システムを指す。

ことが分かる。これは、TackPad の利用者は日本語話者が多いことが理由として考えられる。また、多言語問診票作成システムの用例対訳はベトナム語を除いてほぼ同数の用例数であることが分かる。

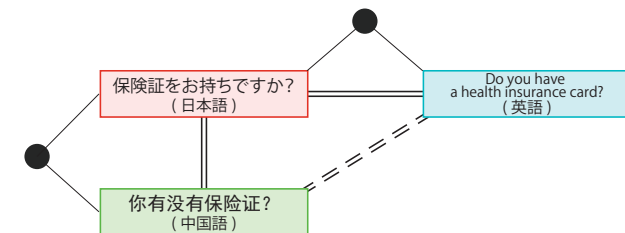
メタノードに結合された用例数を表 2 に示す。TackPad の用例対訳は、3 つもしくは 4 つの用例が結合したメタノードが少ないという結果になった。これは、用例の新規作成を伴わない用例間リンク作成機能の提供を、現時点で行っていないためであると考えられる。また、5 つの用例が結合したメタノードが一番多いという結果になった。これは、収集済みの用例のうち、1 万件あまりの用例はあらかじめ用意した用例対訳をデータベースに直接挿入していることが影響していると考えられる。多言語問診票作成システムの用例対訳は、あらかじめ用例間リンクを多く配置していたため、メタノードに結合された用例数が増える傾

表 3 用例ごとのメタノードとの一致度

Table 3 Degree of agreement to meta-node in example sentences.

一致度	用例数	
	TackPad	問診票
80%	0	4
75%	4	12
67%	4	4
60%	0	4
50%	13,551	11
40%	0	11
33%	27	2
30%	4	4
20%	6	7
合計	13,596	59

・表中の「問診票」は多言語問診票作成システムを指す。



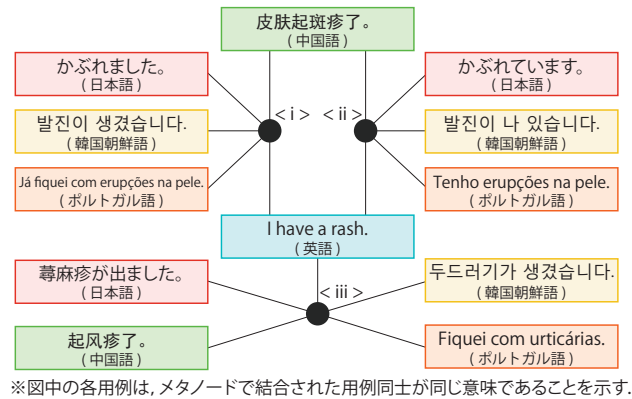
※図中の二重実線は作成済みの用例間リンクを、二重破線は不足用例間リンクを示す。

図 5 不足用例間リンクの例

Fig. 5 Example of an insufficient parallel-text link.

向にあった。

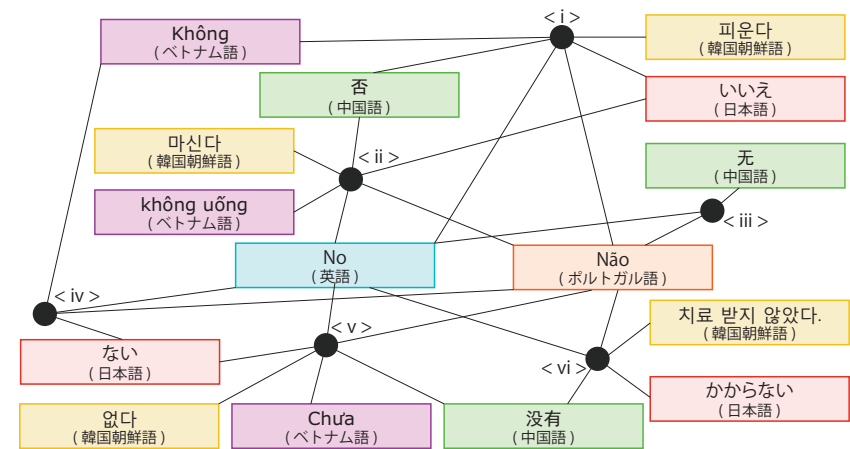
用例とメタノードとの一致度を表 3 に示す。表 3 の用例とメタノードとの一致度をみると、TackPad の用例対訳は、多くの用例に用例間リンクが足りていない可能性が考えられる。用例の新規作成を伴わない用例間リンク作成機能の提供を行い、さらに不足用例間リンクの情報を利用者に提示することで、必要な用例間リンクの増加やメタノードに含まれる用例数の増加が行われると考えられる。なお、表 3 の値は、用例とメタノードごとに計算しており、実際に必要な用例間リンクの数を求めているわけではないため、実際に必要な用例間リンクの数は表 3 の値よりも少なくなる。また、表 3 より、TackPad の用例対訳は一致度が 50% となっているものが多かったことが分かる。一致度が 50% となっていた用例対訳の



※図中の各用例は、メタノードで結合された用例同士が同じ意味であることを示す。
 図 6 TackPad から抽出されたメタノードの例
 Fig. 6 Example of extracted meta-node from TackPad.

例を図 5 に示す。TackPad で一致度が 50%となっているものは図 5 のような形であった。図 5 では、英語と中国語の用例を用例間リンクで結合すると新たに 3つの用例を持つメタノードを作成できることが、本手法の適用により可視化できる。また、多言語問診票作成システムの用例とメタノードとの一致度をみると、少ないながらも用例間リンクが足りていない可能性が見られる。特に、一致度が 80%(4/5 一致) や 75%(3/4 一致) など、用例間リンクによる結合の可能性が高い用例を発見できていることが分かる。

作成されたメタノードの一例を図 6 と図 7 に示す。図中の各用例は、メタノードで結合された用例同士が同じ意味であることを示す。なお、用例間リンクは省略している。図 6 の用例対訳群は、TackPad から抽出されたもので、英語 1 用例、中国語 2 用例、日本語、韓国朝鮮語、ポルトガル語が各 3 用例が結合されていた。図 6 のメタノード <iii> には病名が含まれた用例が結合されている。また、メタノード <i>, <ii> にはニュアンスが多少異なる症状の用例が結合されており、メタノードによる分類を行えていることが分かる。図 7 の用例対訳群は多言語問診票作成システムから抽出されたもので、英語、ポルトガル語が各 1 用例、日本語、中国語、ベトナム語が各 3 用例、韓国朝鮮語が 4 用例結合されていた。図 7 のメタノード <i>, <ii>, <v> は 6 言語の用例がリンクされているが、メタノード <iii>, <iv>, <vi> はそれぞれ 3 言語、4 言語、5 言語の用例のみしかリンクされていないことがわかる。これらの用例対訳に対しては、3.3 節で述べた不足用例間リンクの発見手法を用いると、メタノードに含まれる用例の数が増えると考えられる。



※図中の各用例は、メタノードで結合された用例同士が同じ意味であることを示す。
 図 7 多言語問診票作成システムから抽出されたメタノードの例
 Fig. 7 Example of extracted meta-node from multilingual interview-sheet composition system.

なお、本稿で提案した不足用例間リンクの発見手法は、調査対象となる用例数が増えると計算時間が増える恐れがある。しかし、不足用例間リンク発見手法の適用は、用例間リンクが作成されたときのみ行う。このため、一度に行う処理も少なくなることから、不足用例間リンク発見手法の処理時間は実運用では問題にならないと考えられる。

5. おわりに

本稿では、一対多、多対多の関係にある用例対訳グラフを提供可能な用例対訳とするために、メタデータを利用したデータ構造を提案した。また、手法運用時の問題解決として、新たな用例間リンクの発見手法や、メタノードに意味を持たせるための方法について提案、実装を行った。今後は、実システム上での継続的な運用と、メタノードで結合された用例対訳の提供を行う。

謝辞 本研究の一部は、科研費基盤研究 (B)(22300044) の助成を受けたものである。

参 考 文 献

- 1) Takano, Y. and Noda, A.: A temporary decline of thinking ability during foreign language processing, *Journal of Cross-Cultural Psychology*, Vol.24, pp.445-462 (1993).
- 2) Aiken, M., Hwang, C., Paolillo, J. and Lu, L.: A group decision support system for the Asian Pacific rim, *Journal of International Information Management*, Vol.3, No.2, pp.1-13 (1994).
- 3) Kim, K.J. and Bonk, C.J.: Cross-Cultural Comparisons of Online Collaboration, *Journal of Computer Mediated Communication*, Vol.8, No.1 (2002).
- 4) Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration, *IEEE/IPSI Symposium on Applications and the Internet (SAINT-06)*, pp.96-100 (2006).
- 5) Sakai, S., Gotou, M., Tanaka, M., Inaba, R., Murakami, Y., Yoshino, T., Hayashi, Y., Kitamura, Y., YumikoMori6, T.T., Naya, Y., Shigeno, A., Matsubara, S. and Ishida, T.: Language Grid Association: Action Research on Supporting the Multicultural Society, *International Conference on Informatics Education and Research for Knowledge-Circulating Society (ICKS-08)*, pp.55-60 (2008).
- 6) 高嶋愛里：在日外国人支援活動：京都における「医療通訳システムモデル事業」, 国際保健支援会 2 (2005).
- 7) 宮部真衣, 吉野 孝, 重野亜久里：外国人患者のための用例対訳を用いた多言語医療受付支援システムの構築, 電子情報通信学会論文誌, Vol. J92-D, No. 6, pp.708-718 (2009).
- 8) 福島 拓, 吉野 孝, 重野亜久里：用例対訳を用いた多言語問診票作成システムの開発と評価, 情報処理学会研究報告, グループウェアとネットワークサービス研究会, Vol.2011-GN-78, No.14, pp.1-7 (2011).
- 9) Chen, J., Chau, R. and Yeh, C.-H.: Discovering Parallel Text from the World Wide Web, *ACSW Frontiers'04: Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation*, Vol.32, pp.157-161 (2004).
- 10) Tanaka, Y.: Compilation of a multilingual parallel corpus, *Proceedings of PACLING 2001*, pp.265-268 (2001).
- 11) Bond, F., Nichols, E., Appling, D.S. and Paul, M.: Improving Statistical Machine Translation by Paraphrasing the Training Data, *Proceedings of IWSLT 2008*, pp. 150-157 (2008).
- 12) 福島 拓, 宮部真衣, 吉野 孝, 重野亜久里：医療分野を対象とした多言語用例対訳収集 Web システム TackPad の開発, マルチメディア, 分散, 協調とモバイル (DI-COMO2008) シンポジウム, pp.1030-1036 (2008).
- 13) 梶 博行, 森本康嗣, 相菌敏子, 山崎紀之, 飯田恵子, 内田安彦：コーパス対応の関連シソーラスナビゲーション, 情報処理学会研究報告, データベース・システム研究会, Vol.1999-DBS-118, pp.97-104 (1999).
- 14) 李 航：ハイパー・シソーラスとその学習, 情報処理学会研究報告, 自然言語処理研究会, Vol.1992-NL-92, pp.81-88 (1992).
- 15) 福井健司, 柏岡秀紀：対訳文選択のための用例翻訳用シソーラスの構築, 情報処理学会研究報告, 自然言語処理研究会, Vol.2006-NL-124, pp.47-54 (2006).
- 16) 小島秀一, 高須淳宏, 安達 淳：Web ページ群の構造解析とグループ化, *NII journal*, Vol.4, pp.23-35 (2002).
- 17) 岡田直樹, 谷川恭平, 土方嘉徳, 西田正吾：交グラフと意味的解析を利用したコミュニティ発見手法, Web とデータベースに関するフォーラム (WebDB Forum 2010) (2010).