

## 英会話用知的 CAI システムの教材作成支援における文生成の利用

2 T-7

瀬戸 美枝<sup>†</sup> 浅野 雅代<sup>†</sup> 宮崎 敏彦<sup>††</sup> 山本 秀樹<sup>††</sup>  
<sup>†</sup>(株) 沖テクノシステムズ ラボラトリー <sup>††</sup>沖電気工業(株)

## 1 はじめに

著者らは、自然言語を使った自由な対話をシミュレートすることにより学習者の会話力を高めることを目標としたシステム [1, 2] の研究開発を行なっている。

本論文では、本システムの教材作成支援における文生成 [3] の利用について述べる。

## 2 英会話用知的 CAI システムの教材作成

本システムの教材作成は、

1. 会話のストーリーを決める
2. 会話のシナリオを作成する
3. シナリオに出てくる文の言い換えを作る
4. 3の文に対して、学習者の誤りを予想し、誤りを含んだ文を作成する
5. 3,4の文を教材のデータ形式に変換する

といった手順で行なわれる。2,3については、システムの発話文及び学習者の発話文を対象とし、4については、学習者の発話文のみを対象とする。4は、学習者の誤りを発見するために使用するバグーモナルである。教材作成者にとって、3~5の作業は、1,2の作業結果をもとに行なう作業であり、1,2の作業に比べると創造的な面が少ない作業であるといえる。また、より充実した教材を作成するには、3,4で多様な文を作成することが重要であるが、それは教材作成者にとって多大な労力を要するものである。本システムの実用化を図るには、いかに少ない労力でこれらの文を用意することができるかが重要な要素となる。

そこで、2を入力として、それを解析し、言い換え文や誤りを含んだ文を生成するための知識を使って、3,4を生成するシステム(教材作成支援のための文生成システム: Sentence Generation Supporting System 以下, SGSS と呼ぶ)について提案する。

## 3 文生成を利用した教材作成支援システム

## 3.1 システムの構成

SGSSの構成を図1に示す。SGSSは、解析部と生成部のモジュールと、会話の状況を蓄積する状況メモリから構成される。

## 3.2 意味表現と状況メモリ

解析部は、シナリオのある一文を入力とし、入力文に対して構文解析を行ない構文木を生成する。次に意味解析を行ない意味表現

を生成する。さらに談話理解を行ない、会話の状況を状況メモリに蓄積する。

生成部は、解析部で生成されたある一文の構文木と意味表現を入力とし、状況メモリを参照して、表1に示すような様々なタイプの文 [4] を生成する。

本システムの意味表現は、属性名と属性値の対の集合である格フレームの形式である。属性値の属性は、再帰的に格フレームの形式で表すことができる。「友人に電話をかけて食事のアポイントをとる」という内容のシナリオのある一文 "This is Taro Tanaka of XYZ." に対する意味表現の例を、表2に示す。

状況メモリは、意味表現と同様にフレーム形式で表現される。状況メモリには、会話の状況(表3)、学習者とシステムの発話履歴などが格納されている。

## 4 実現方式

本章では、解析部で得られた "This is Taro Tanaka of XYZ." の構文木(図2)と意味表現(表2)を入力とし、状況メモリ(表3)を参照して、表1で示したタイプの文を生成する方法の一例について述べる。その結果生成される文の例を表4に示す。

## 4.1 正しい言い換え文の生成

状況メモリ(表3)の scene, 意味表現(表2)の function には、それぞれ「電話」、「名乗る」が格納されているので、ここでは「電話で名乗る」という場面における論理的、文法的に正しい文の生成が要求される。入力文では、caller(呼び出し人)の名前(Taro Tanaka)と社名(XYZ)を一文で表現しているが、この二つの情報を二文に分け、表4に示す PARA2-1 のように表現することも可能である。このとき生成部は、図2で示した "This is Taro Tanaka of XYZ." の構文木より "of XYZ" を抜いた構文木から第一文を、図3に示す構文木から第二文を生成する。

また、決まり文句に対応するには、テンプレートを用いて文を生成する場合もある。

## 4.2 論理的に問題のある文の生成

生成部は、図2で示した構文木の "this", "is" をそれぞれ "I", "am" に置き換えることにより、「電話で名乗る」という場面では相応しくない文 (UN-US) を生成する。

## 4.3 文法的に問題のある文の生成

生成部は、図2で示した構文木の \*1 で示される np を木のルートとして、主語と述語が欠落した文法的に不完全な文 (G-INC) を生成する。

ここでは、爆発的に文を生成することを防ぐために、英語を母国語としない者が犯しやすい誤り [4] を優先している。

An Application of Generating Sentences to Authoring System of an ICAI System for Conversational English.

<sup>†</sup>Mie SETO, Masayo ASANO

<sup>††</sup>Toshihiko MIYAZAKI, Hideki YAMAMOTO

<sup>†</sup>Oki Technosystems Laboratory, Inc

<sup>††</sup>Oki Electric Industry Co., Ltd.

5 おわりに

本論文では、英会話のシミュレーションを行なう知的CAIシステムの教材作成支援における文生成の利用について述べた。文生成を利用することにより、少ない労力で多様な文を作ることができる。また、生成された様々な文は、学習者の入力文の認識率向上や誤り指摘、あるいはシステム側の発話の多様化に有用である。

参考文献

- [1] 山本ほか: 会話シミュレーションを基にした語学訓練用知的CAIシステムの構成, 情報処理学会論文誌, Vol. 30, No. 7, pp. 908-917 (1989).
- [2] 山本ほか: 音声入力可能な英会話用知的CAIシステム, 電子情報通信学会教育工学研究会, ET-91-107(1991).
- [3] 浅野ほか: Communicative Approachに基づく英会話ICAIシステムの文生成について, 情報処理学会自然言語処理研究会報告, 81-10(1991).
- [4] 垣田直巳(監修): 英語の誤答分析, 大修館書店(1983).

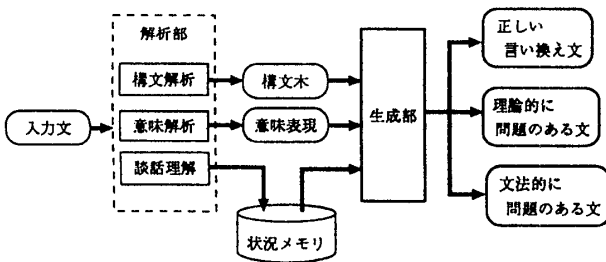


図1: SGSSの構成図

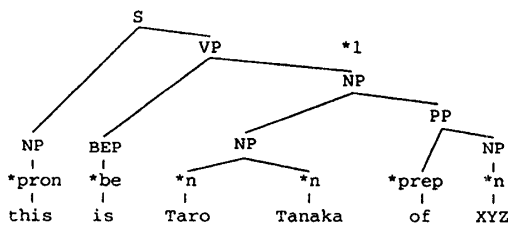


図2: "This is Taro Tanaka of XYZ."の構文木

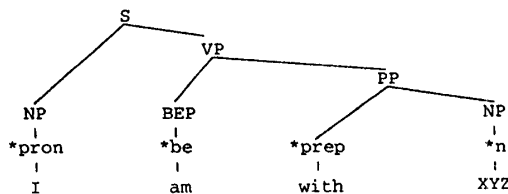


図3: "I am with XYZ."の構文木

表1: 生成する文タイプとその内容

文タイプ	内容	
正しい言い換え文	PARA1	高い難易度の文
	PARA2	標準的難易度の文
	ARA3	低い難易度の文
論理的に問題のある文	L-INC	論理的に不完全な文: 必要な情報が欠落している文
	L-ERR	論理的に誤った文: 表現が類似して意味が異なる文
	UN-US	語の用法が不適切な文: その場に相応しくない表現の文
文法的に問題のある文	G-INC	文法的に不完全な文: 主語, 述語, 目的語, 補語などが欠落している文
	G-ERR	文法的に誤った文: 文の構造や語形変化などが誤っている文

表2: 意味表現の例

属性名	属性値
object (対象)	caller (呼び出し人)
function (言語機能)	introduce-oneself (名乗る)
.....	.....

表3: 状況メモリの例

属性名	属性値	下位属性	
		属性名	属性値
scene (場面)	telephone (電話)	first name	taro
caller	human (人)	last name	tanaka
.....	.....	company	XYZ
.....	.....	.....	.....

表4: 生成文の例

文タイプ	No	生成文
PARA2	1	This is Taro Tanaka. I am with XYZ.
	2	This is Taro Tanaka speaking. I am with XYZ.
	3	Taro Tanaka speaking. I am with XYZ.
L-INC	1	This is Taro Tanaka.
L-ERR	1	This is Taro Tanaka of the XYZ.
UN-US	1	I am Taro Tanaka of XYZ.
	2	My name is Taro Tanaka. I am with XYZ.
G-INC	1	Taro Tanaka of XYZ.
G-ERR	1	This is XYZ's Taro Tanaka.