

マルチメディア理解システム IMAGES-M における 音声に関する意味処理について

天野 幹郎 村田 聖 横田 将生
福岡工業大学

1. はじめに

我々は、音声・文字・静止図形・動画像など種々のメディアを統合的に理解し、相互参照を可能とするマルチメディア理解システム IMAGES-M の実現を目指している。

人間が日常用いる言語は大別して文字言語と音声言語の2通りある。音声言語は文字言語情報(内容)だけでなく話者個人の情報や情緒性情報などを含んでおり、文字言語よりも効率よく情報を表現する機能を有している。

本稿では、IMAGES-M の音声処理部における音声合成処理を意味処理の立場から、陽に表現されないことの多い(願望などを含む広義の)信念を相手に伝達するため音声に含まれる情報について述べる。

2. 情報体系と情報媒体

情報は、媒体によって表現・伝達される意味内容(すなわち概念)でありセマンティクスの取り扱い対象のことである。ここでは心像(イメージ)に帰着される。

情報体系は図1に示すような1)対象世界(W), 2)概念体系(C), 3)表現体系(R)およびそれらに係わる処理体系である。W および C の関係を明確に受けて通達先に表現するものが R である。

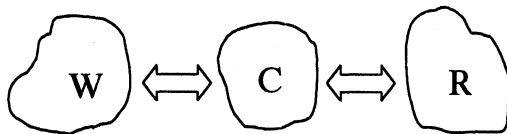


図1 情報体系の概念図

情報媒体は、ある情報体系が情報を他の情報体系に伝達するための表現手段である。ここでは、情報体系に共通の syntax と semantics を持った言語が主となり、そうでない画像や音声は補助的に働く多重媒体(multimedia)を想定している。このような媒体は、言語を核として図1に示す体系(すなわち表現体系(R))をなしている。

3. コミュニケーション

コミュニケーションは、情報体系の相互作用と定義できる^[1]。ここでの相互作用は、種々の媒体(言語、画像、音声)を介して情報(知識、信念)を交換することにより共通の部分体系を形成する事である。

情報の最小単位をメッセージ(message:通達)と呼ぶことにすると、コミュニケーション中におけるメッセージ(m)は式のようなになる。

$$m = (s, h, b)$$

s: 話し手(speaker)

h: 聴き手(hearer)

b: 信念 (belief)

$$b = (e, r)$$

e: 概念(対象世界における事象の情報)
(event)

r: 依頼(実現するとより幸せになるという信念)
(request)

人間は、依頼を受けた相手がどうかにより反応は異なるが、好ましいならば喜ぶことをするという傾向がみられる。表1の下線部分が受け手の期待される反応である。

表 1 文の種類と文の意味表現

文の種類： 例文	e の記述	R の内容	R の記述 $r \in B_s$ b-val：信念値という属性
平叙文： 今日は晴れです	$(x,y,晴れ,晴れ,質)$ $\wedge 天気(y)$	h が e を信ぜよ。	$(f \supset_3(f,s,p,happier,emo))$ $\wedge f=(h,e/.1,b-val,h)$
疑問文： 「今、何時ですか？」	$(\lambda q)(\phi,t,q,q,time-point,k0)$ $\Pi(\phi,t,現在,現在 time-point,k1)$ $\wedge time(t)=(\lambda q)e(q)$	真偽を判定せよ。あるいは、真となるものを答えよ。	$(f \supset_3(f,s,p,happier,emo))$ $\wedge f=(h,e(\alpha),h,s,inf-loc)$ $\wedge (h,e(\alpha),1,1,b-val,h)$ < α は定項>
「今、10 時ですか？」	$(\lambda q)(\phi,t,10,10,time-point,k0)$ $\Pi(\phi,t,現在,現在 time-point,k1)$ $\wedge time(t)$ k0：時計による物理的時刻 k1：現在・過去・未来という時刻		$(f \supset_3(f,s,p,happier,emo))$ $\wedge f=(h,e',h,s,inf-loc)$ $\wedge e'=(h,e,q,q,b-val,h)$ <知らない>という q 値も存在する
命令文： 「新聞を持って来い」	$(\phi,h,p,s,posi) \Pi(h,x,p,s,posi)$ $\wedge paper(x)$	e を実現せよ (真とせよ)。	$(f \supset_3(f,s,p,happier,emo))$ $\Pi(\phi,s,優位,優位,emo.h)$ $\wedge f=(h,e/.p,place)$ <h が e を起こす> <精神的に優位> f は事実でないといけない
感嘆文： 「何と速い車でしょう！」	$(\phi,x,very.fast,very.fast,速度)$ $\wedge car(x)$ $\wedge (x,s,感動,感動,emotion)$	h が e を信ぜよ。	$(f \supset_3(f,s,p,happier,emo))$ $\wedge f=(h,e/.1,b-val,h)$

4 まとめ

今回、心像意味論に基づき音声におけるコミュニケーションの概念化と意味表現記述について試みた。今後は、心のモデル⁴などとの連携を行い、滑らかで自然な応答のできるシステムとして目標達成を図りたい。

参考文献

- [1]横田将生：心像意味論に基づくコミュニケーション論，福岡工業大学情報工学研究所集報 第5巻(1994)
- [2]横田将生，天野幹郎：自然言語理解を目的とする心の計算モデルについて，情報処理学会研究報告 (2001・NL-141-7)