

表現メディアとしてのデジタルヒューマンと その環境依存発話に関する検討

亀田弘之¹・飯田仁²

和文抄録 サイバー空間内にデジタルヒューマンを表現メディアとして配置することにより、従来のものとは異なる新たなヒューマン・コンピュータ・インタフェースを実現する構想とその実現例に関する検討結果について報告する。具体的には、「メディアとしてのヒト」およびその具現例である「デジタルヒューマン」について基礎的検討を行うとともに、デジタルヒューマンに対して、環境依存型発話機能を持たせるための方法とそのプロトタイプ実現例について述べ、環境依存型発話の基本的必要性・有効性を確認した。

A Note on Digital Human as Media for Expressions and its Environment-Dependent Speaking

Hiroyuki KAMEDA¹ and Hitoshi IIDA²

Abstract This paper reports our project to realize a new type of human-computer interfaces and also describes its implementation of a prototypical system, on which digital humans are configured as media to express contents in cyber spaces. We also describe our considerations of human beings as media, digital humans as an instances of them, and how to realize digital humans' environment-dependant speaking. An example of implementation of environment-dependant speaking is also presented in the case of weather forecasts.

1. はじめに

周知のように、ブロードバンドあるいは無線 LAN の整備によるインターネットの大衆化、さらには、携帯電話の普及・高機能化により、現代社会は日々ますます高度に情報化され国際化されつつある。いまや書籍・絵画・音楽などをはじめとする多くの情報がデジタル化されており、我々はそれらをネットワーク経由でいとも容易に手に入れることができる。一方、このような状況であるにもかかわらず、如何に多くの情報が手元にあると真に必要とされる情報を迅速かつ適切に入手することは、今日でもなお困難である。

この問題点を解決するために、検索エンジンの高度化、さらには情報呈示などの情報デザインの観点からの改善の試みが積極的になされている。その結果以前に比べて多くの改善がなされたが、これら従来型の改良にはやはり限界があるといわざるを得ない。

このような状況に鑑み、本研究では従来のようなタイプのインタフェースではなく、「人間型インタフェース」を新たに提案する。具体的には、コンテンツあるいはコンテンツ群としてのサイトに関する知識（メタ知識も含む）を有するとともに、人間（例えばサーチャーや図書館司書）と同様の振る舞いをするのできる「デジタルヒューマン」（以下、「DH」とも記す）を提案する。

1: 東京工科大学コンピュータサイエンス学部・
School of Computer Science, Tokyo Univ. of Tech.

2: 東京工科大学メディア学部・School of Media Science,
Tokyo Univ. of Tech.

2. デジタルヒューマンとその役割

人間が最もコミュニケーションをしている相手は、人間が社会的な動物であることから明らかのように、「人間」である。つまり、人間にとってもっとも慣れ親しみ、かつ、人間相互の意思疎通をもっとも自然・的確に実現できるのは「人間」という表現メディアである。このことは、サイバー空間においては、その中のバーチャルなヒューマン、すなわち、DHこそが我々人間にとって自然で適切なコミュニケーション用インタフェースであることを示唆している。

このような観点から、以下にまず人間(ヒト)の各部位の意思疎通における機能を重要なものに関して簡単に述べ、その後、DHのサイバー空間における役割について述べる。

2.1 コミュニケーションにおけるヒトの各部位の役割

- (1) 口：発声器官の一部であり、音声などの音を発する機能がある。音声や泣き声などにより、言語的な意味情報のみならず感情・体調などの情報も伝達することができる。
- (2) 耳：聴覚器官であり、音声を聞き取る機能をもつ。例えば、「耳を澄ます態度」を示すことにより、「積極的に聞き取るうとする姿勢」を示せたり認識したりすることができる。
- (3) 顔：口・目・眉毛・頬などから構成されており、これらの配置を微妙に変化させることにより、気持ちを表現することができ、また、それらを検知することができる。
- (4) 手：本来は物をつかむための器官であるが、意思疎通の際には、ボディランゲージ(body languages)のための器官として使用される。これは文化に依存するがコミュニケーションには有用なものである。さらに、これが体系化・洗練されたのが手話(sign languages)であるとも考えられる。なお、手は文字言語のためにも、すなわち、文字を記述するためにも使用される。

社会的存在である人間(ヒト)は、このように身体の各部位をコミュニケーションのために利用している。従って、サイバー空間と実世界とのインターフェースは「人間型」であることが望まし

いことは明らかであろう。以下、人間型インタフェースとしてのDHについてさらに詳しく述べる。

2.2 デジタルヒューマンとその役割

例えば、現在の Web での情報提供方式は、ウェブページ相互にリンクされたハイパーテキスト形式であり、これらを検索エンジンで検索する方式である。これに対して、本研究で提案する DH を用いる方式では、「コンテンツは従来と同じでよいが、コンテンツの内容(記述内容)およびコンテンツの構成に関する知識(メタ知識)を DH が有しており、その DH があたかも人間がサイバー空間に情報案内人として配置されているかのように振舞う」というものである。以上をまとめると、DHの主な役割は以下ようになる。

- (1) **司書機能の提供:** DHは、コンテンツにどのような情報・知識が記載されているかを知っており、例えば、ユーザに「XXXという機能を持った新しい製品があるか知りたいんですが?」と問われたときには、「そのような情報はこのサイトには記載されておりません。こちらのサイトをご覧ください。」と答えて、Web ページ内に記載されているクリックすべき URL を手で指し示したり、あるいは、「これがそうです。」といて当該ページを表示したりする。この機能により DH は表現メディアとみなせる。
- (2) **音声対話機能の提供:** ユーザとの言語によるやり取りは、文字言語(テキスト)による方法(筆談)もありえるが、効率を考えると、音声による対話機能が实际的であろう。
- (3) **意図の明確化の支援:** ユーザと内容的に適切な対話を進めるとともに、ユーザに対してストレスのない自然で円滑な対話を実現するためには意図推定・理解が必要である。さらには、ユーザ自信が発話意図を明確に意識・理解していない場合にも、ユーザの意図を対話により明確にするという支援機能があることが望ましい。

3. 関連研究

DHに関する研究としては、バーチャルヒューマン(Virtual Human、以下、「VH」とも記す)と知的エージェントとがあるが、多くのVHは知

的エージェントの機能を備え持っているので、以下 VH について述べる。

現在、VH としては例えば以下のようなものがある。

- (1) **MS Agent:** Windows のインターフェース機能を高めるための animated character agent。MS Word の OFFICE アシスタンスのバージョンアップされたものに相当し、音声認識機能によるコマンド実行や text-to-speech 型音声合成による対話が可能。
- (2) **パラブラ** ネットコミュニケーション用コミュニケーションエージェント。アバターとしてパーソナルコンピュータ（以下、「PC」と記す。）上に配置し、他のユーザとのチャットやメッセージ交換、さらには、ウィルスチェックやメールチェックの機能を実現している。
- (3) **NetPeople:** 3D キャラクタ(人間)が Web ページとともに現れ、そのページに関する問い合わせに答えたり、必要に応じて、「ここではここをクリックしてください。」などのアドバイスをする。対話は作り込みにより相当な程度対応可能であるとの報告もある。
- (4) **傀儡システム:** 学術的観点から研究・作成されているシステムであり、音声認識技術と自然言語処理技術とを融合させ、音声言語理解と CG によるソフトウェアロボットの実現とその行動制御機能の実現を目指している。関連分野の技術のブレークスルーを目指している。

以上の各 VH のうち、(1)は OS やソフトウェア操作の Help 機能の高度化を目指したものであり、(2)はコミュニケーション支援と知的エージェント機能の提供を目的としたものであり、(4)は関連技術の格段の進歩を目指した学術的研究である。(3)は本研究の目的に近いものではあるが、アニメーション技術としては優れているが、エージェント機能にはまだ問題がある。本研究で提案する DH は、(3)の高度化であるとともに、従来のコンテンツ提供 & 検索の枠組みを根本から変えるものであるところが特徴である。

4. デジタルヒューマンの基本機能

すでに述べたように、デジタルヒューマンとはサイバー空間における仮想的人間であり、そのため、入出力機能としては先にも述べたような人間と同様の機能を自然に備え持つ。具体的には、身体(手・足・顔)を持つとともに、それに付随して目(視覚機能)・耳(聴覚機能)・皮膚(触覚機能)を、さらには内部的には、喜怒哀楽等の情緒の表出・理解、発話理解・生成、推論・知識獲得等の機能(心のモデル)を持つ。

本研究ではそのうち、表現メディアとしてのデジタルヒューマンの基本機能として、まずは以下のものに着目する。

(1) **身体動作:** 場所移動(表示画面内移動・サイバー空間内移動)、身振り(身体表現・手話)であり、表現メディアとしての DH に対して表現の豊かさをもたらす。文化依存的な側面も多く、実用の際には文化人類学的あるいは行動動物学的観点からの注意を払う必要がある。

(2) **表情:** 喜怒哀楽等の情緒的情報の表現とともに、手話などでも付随的に見られるような意思疎通の補助的手段。なお、「表情」は身体動作の一つとも考えられるが、人間における意思疎通では特に重要であるために、別項目とした。なお、表情は先の身体動作ほどは文化依存ではないので、非言語的意思疎通の手段として期待される。

(3) **音声処理:** 音声認識・音声合成機能であり、特にユーザとの対話において重要である。なお、音声言語・文字言語に関わらず、ユーザとの柔軟かつ円滑な対話機能は今後も重要な研究テーマである。DH の成功の可否の 1 要因として、この対話機能があげられよう。

(4) **意図理解:** ユーザの発話目的・意図を推定し、ユーザに違和感を感じさせない対話を実現する。DH が関わるコンテンツの範囲が明確な場合には、当該コンテンツの用途がある程度は把握されているので、ユーザの意図を推定し理解することが容易になる。

上記のもの以外にも、推論・未知語/未知知識処理・矛盾を含む不完全な知識の処理・心的葛藤処理等の機能が考えられるが、本研究ではまず上記の 4 つに焦点をあてて研究することとした。

5. 環境依存発話とその例

DH とユーザが対話をする場合、従来の多くの研究では、ELIZA システムのような単純な対話システムを自然言語処理的観点から格段に高度化させたものの実現が1つの目標となっている。本研究ではそれとともに、ユーザとDHが両者の存在する物理的環境（例えば、天候状況・ネットワークのトラフィック状況など）にも依存する発話の実現を目指す。以下、環境依存発話の例とそれに基づき環境依存発話の有効性について述べる。

5-1. 環境依存発話の例

本研究で提案する“環境依存発話”とは、例えば以下のようなものである。

・**状況設定(気象状況依存発話)**: ある晴れた夏の日。気温も湿度も高く過ごしにくい。ユーザは就職活動のためとある会社を訪問した。その後近所でお昼を食べたいのだがどこで何が食べられるのかが分からない。

[対話例1]

U: この近所でお昼を食べたいんだけど...

DH: おなかいっぱい食べますか?

U: 今日は暑すぎるので、軽くでいいかな?

DH: それでは、お蕎麦はどうですか? すぐ近くに美味しいお蕎麦屋さんがありますよ。

U: 何を食べようかなあ?

DH: 今日は暑いから、冷たいお蕎麦や冷やし素麺なんかがいいのでは?

この対話で、アンダーラインを付けた部分が、本研究で目指している環境依存の応答例である。次の例も同様である。

[対話例2]

U: いつものサイトにつながらないんだけど。

DH: ちょっと待ってください。

U: 接続先のコンピュータが止まっているのかなあ?

DH: ネットワークが込んでいるみたいですね。30分後にもう一度アクセスしてみてください。

5.2 環境依存発話の有効性

上記の例からも察せられるように、環境に依存する発話ができることで、単なるシナリオ通りの発話・対話ではなく、表現・発話内容に広がりや深みを持たせることができ、より自然な対話を実現させることが可能となる。このことは、Human-Computer Interface の高度化に大いに寄与するものと考えられる。

6. 対話における状況の分類

6.1 状況の分類(試案)

本研究ではまず DH の立場に視点を限定して、対話における状況を以下のように外部状況と内部状況との2つに大別・分類した。

(1) 外部状況:

- ・時空的状況(時空点=(x, y, z, t))
- ・外界の環境状況(気象状況など) など

(2) 内部状況:

- ・心的状況(感情 emotion・気分 mood)
- ・知識:
 - ・固定的知識(static knowledge)
 - ・体系化知識: 世界に関する知識(organized knowledge)
 - ・非体系化知識: 常識(ad hoc knowledge)
 - ・一時的知識(temporal knowledge)
 - ・自己の状況に関する知識
 - ・直截的内観に基づく自己の知識
 - ・他者から見た自己に関する知識
 - ・他者の状況に関する知識 など

本研究における状況記述の枠組みは、まず外部状況を中心に考えることとした。

6-2. 環境依存発話機能の実現例

以下では上記の分類を踏まえて、気象状況という環境に依存する発話を例として取り上げ、環境依存発話の実現例(プロトタイプ)について述べる。具体的には以下のような設定を考える。

(1) 対象: 天気に関する文(ローカルな地域の気象状況)

(2) 状況(環境)の記述(一般形)と状況と文との対応関係: 気温や湿度あるいは晴れか雨かなどの(狭義の)天気状況を表すためのパラメータ

として、状況パラメータを考え、状況パラメータの値の組を状況ベクトル、さらに、状況ベクトルの集合を状況ベクトル空間 E と呼び、E から文の集合への写像を文生成関数 F と呼ぶこととする。

状況パラメータ群： p_1, p_2, \dots, p_n

状況ベクトル： $e_i = (p_{1i}, p_{2i}, \dots, p_{ni})$

状況ベクトル空間 $E = \{e_1, e_2, \dots, e_m\}$

文の集合： $L = \{s_1, s_2, \dots, s_k\}$

文生成関数 F： $E \rightarrow L$

(3) 気象状況文発話への適用例： 上記の(2)に基づき、例えば以下のような設定を設ければよい。まず、状況パラメータとして、天気パラメータ p_1 、気温パラメータ p_2 、湿度パラメータ p_3 を設け、それぞれ以下のようにすることができる。

- ・天気パラメータ p_1 { 晴れ, 曇り, 雨 }
- ・気温パラメータ p_2 [] { 低, 中, 高 }
- ・湿度パラメータ p_3 [%] { 低, 中, 高 }

これらのパラメータがさまざまな値を持つときに、それに応じて以下のような文のいずれかを生成させればよい。

- ・文集 $L = \{$ 今日は晴れです。今日は曇りで今日は雨です。今日は暑いです。今日はとても暑いです。今日は涼しいです。今日はとても涼しいです。今日は寒いです。今日はとても寒いです。今日は蒸し暑いです。今日はとても蒸し暑いです。 $\}$

以上のことを前提として後は文生成関数 F を決定すればよい。文生成関数 F は原理的には単に対応表を網羅的に記述すればいいのだが、一般にはそのような方式を採用するのは L が小規模でない限り得策ではない。本研究では Chomsky の教えを模倣して、次のような形式文法の枠組みを採用した。なお以下では、生成文に日時の成分を含ませるために、さらに日時パラメータ p_0 を新たに導入した。

言語 $L=L(G)$

文法 $G = (V_n, V_t, \text{ , PR})$

ただし、

- ・非終端記号の集合 $V_n = \{$ 日時成分, 天気分, 晴天, 曇天, 雨天, 句点 $\}$

・終端記号の集合 $V_t = \{$ 今日, 明日, 朝, ... $\}$

PR = {

日時成分 + 天気成分 + 句点

日時成分 今日は(p_0 =今日) |

明日は(p_0 =明日) | 朝は(p_0 =朝) |

朝のうちは(p_0 =朝) |

昼間は(p_0 =昼) | 夜は(p_0 =夜) |

夜間は(p_0 =夜)

天気成分 晴天(p_1 =晴) | 曇天(p_1 =曇) |

雨天(p_1 =雨) | 荒天(p_1 =荒)

晴天 晴れです(p_1 =晴) |

いい天気です(p_1 =晴) |

穏やかないい天気です(p_1 =晴, p_2 =中) |

からっとしてとてもいい天気です(p_1 =晴,

p_3 =低) | 晴れます(p_1 =晴) |

晴天です(p_1 =晴)

曇天 曇りです(p_1 =曇) |

まあまあ天気です(p_1 =曇) |

曇ります(p_1 =曇) |

曇天です(p_1 =曇)

雨天 雨です(p_1 =雨) |

あいにくの天気です(p_1 =雨) |

雨が降ります(p_1 =雨) |

雨天です(p_1 =雨)

荒天 台風です(p_1 =荒) |

大荒れです(p_1 =荒) |

ひどい天気です(p_1 =荒)

句点 。 }

(4) 気象状況文発話への適用結果： 上記(3)の設定により、例えば、 $e = \{p_0$ =今日, p_1 =晴, p_2 =中, p_3 =低, $\}$ の状況のとき、文生成関数はこの環境に依存して、「今日は穏やかないい天気です」という文を生成する。

(5) インプリメント例： 上記(4)は次のように形式的に Prolog 言語でインプリメントすることができる。

%状況(環境)設定に応じて生成される文が異なる。

r1(A) :- s(A,状況([[p0=明日],[p1=雨],any]]).

r0(A) :- s(A,状況([[p0=_],[p1=_],any]]).

%任意の状況を許容する場合

s(A,X):-s(A,[],X).

s(A,D, 状況([C1,C2,C3])):-日時成分(A,B,C1), 天気成分(B,C,C2), 句点(C,D,C3).

日時成分([今日は|T],T,[p0=今日]).
日時成分([明日は|T],T,[p0=明日]).
日時成分([朝は|T],T,[p0=朝]).
日時成分([朝のうちは|T],T,[p0=朝]).
日時成分([昼間は|T],T,[p0=昼]).
日時成分([夜は|T],T,[p0=夜]).
日時成分([夜間は|T],T,[p0=夜]).

天気成分(A,B,P1):-P1=[p1=晴],晴天(A,B,P1);
P1=[p1=曇],曇天(A,B,P1);
P1=[p1=雨],雨天(A,B,P1);
P1=[p1=荒],荒天(A,B,P1).

晴天([晴れです|T],T,[p1=晴]).
晴天([いい天気です|T],T,[p1=晴]).
晴天([穏やかないい天気です|T],T,[p1=晴,p2=低]).
晴天([からっとしてとてもいい天気です|T],T,
[p1=晴,p3=低]).
晴天([晴れます|T],T,[p1=晴]).
晴天([晴天です|T],T,[p1=晴]).

曇天([曇りです|T],T,[p1=曇]).
曇天([まあまあ天気です|T],T,[p1=曇]).
曇天([曇ります|T],T,[p1=曇]).
曇天([曇天です|T],T,[p1=曇]).

雨天([雨です|T],T,[p1=雨]).
雨天([あいにくの天気です|T],T,[p1=雨]).
雨天([雨が降ります|T],T,[p1=雨]).
雨天([雨天です|T],T,[p1=雨]).

荒天([台風です|T],T,[p1=荒]).
荒天([オ!大荒れです|T],T,[p1=荒]).
荒天([ひどい天気です|T],T,[p1=荒]).

句点([. |T],T,any).

6-3. 考察と問題点

本手法は、言語学的には文脈自由文法の形式を基本としているが、書き換え規則適用のための条件が状況パラメータ（発話環境）の値により記述されているために、記述能力的には文脈自由文法以上のレベルにあると考えられる（参考文献[2]の233ページ参照）。

- 今後は基本的にはこのような枠組みとし、
- ・状況パラメータの種類（社会的状況など）を増やす
 - ・状況パラメータと生成規則との関連をさらに検討する。可能ならばこれらを形式上分離し、状況パラメータを書き換え規則適用のメタ規則化する。
 - ・複数の文（簡単な文章）が生成できるように、生成規則を拡張する。すなわち、対話構造レ

ベルの規則まで拡張する。

- ・「発話者の心のモデル」の導入。ユーザの心情や価値観を踏まえた発話を実現するためには、これが有効であると考えられる。

などの工夫・拡張が必要であろう。

特に最後に言及した「心のモデル」であるが、例えば、「心の状態」を「喜怒哀楽(身体の状態などにより変化)」と定義すると、DHは自分自身に直接的に起因する感情ではなく、相手（語りかける対象者（ユーザ）以下、「U」と記す。）の身体的・社会的等の状況に起因する感情に応じ、DHの心の（内部）状態を変化させる現象を捉えることができる。

例えば、「U が久しぶりの休日をとることができ、明日やっと友人と海水浴に行くことになった（社会的状況）明日は風もなく、湿度もほどほどで、気温も十分高そうないい天気になりそう（自然（気象）状況）」このような状況の下、標準発話「明日は晴れです。」を「明日は晴れでいい天気です。」や、「明日は晴れで、気温も高くいい天気です。（ここで、「いい」とは社会的に「いい」と言う意味）」さらには、「明日は海水浴日和です。」といった発話を生成することが見込める。すなわち、環境依存型発話を実現することができる。

7. おわりに

デジタルヒューマンを表現メディアとして捉えることの有効性と、デジタルヒューマンに環境依存型発話させる意義と方法等について述べた。

謝辞 本研究を推進するにあたり、多くの助言あるいは示唆を与えてくださった横井俊夫・金子満両教授（東京工科大学）とともに、東京工科大学学内共同研究の共同研究者各位に感謝いたします。なお、本研究は東京工科大学学内共同研究の一部として推進されたものである。

参考文献

- [1] 横井・金子・岡本・植木・永田・大野・亀田：Virtual Humanoidに関する調査研究，東京工科大学共同プロジェクト報告書(2002).
- [2] 守屋：形式言語とオートマトン，サイエンス社(2001).