

## コミュニケーション基礎実験計画とその環境設計

山本 吉伸          仁木 和久

電子技術総合研究所

E-mail:yoshinov@etl.go.jp, niki@etl.go.jp

我々の経験では、ネットワークを介して行なうコミュニケーションに比べて、直接対面して行なうコミュニケーションの方が柔軟性に富んでおり、有意義で集中的かつ効率的な会合を持たたとの印象を受け易い。このような印象(メディア不協和感)はコンピュータネットワークの特性に依存する可能性がある。本稿では、このメディア不協和感を実証的に検討するためのプロジェクトを紹介する。本プロジェクトは、ヒューマン・ヒューマンコミュニケーションの理論(モデル)を構築することを目標とする「認知理論研究プロジェクト」と、認知理論研究プロジェクトをはじめ数種の動画音声通信関連研究の実施環境整備を目標とする「マルチメディア環境整備プロジェクト」から構成される。

## A design for communication experiment and environment

Yoshinobu Yamamoto      Kazuhisa Niki

Electrotechnical Laboratory

In our experience, we find direct communication more efficient and flexible than mediated communication (where information between two agents is communicated via a third agent). However, use of a mediating agent is necessary when, for example, there are differences in protocol between the communicating agents. This report introduces our research project which attempts to address the cognitive problems arising from mediated communication. This project is composed of two subprojects: (1) "Cognitive theory research project" - concerned with modeling human-to-human communication; and (2) "MultiMedia environment development project" - concerned with the construction of a multi-media environment on the internet for researchers.

## 1 はじめに

本稿では、我々が進めている動画音声情報交換システム運用実験プロジェクトについて述べる。本プロジェクトは、ヒューマン・ヒューマンコミュニケーションの理論(モデル)を構築することを目標とする「認知理論研究プロジェクト」と、認知理論研究プロジェクトをはじめ数種の動画音声通信関連研究の実施環境整備を目標とする「マルチメディア環境整備プロジェクト」から構成される。最初に、我々の本来の目的である認知理論研究プロジェクトを紹介し、次に認知理論研究プロジェクトの基盤となるマルチメディア環境整備プロジェクトを紹介する。

我々は、マルチメディア環境を広い範囲で相互接続して実験、利用したいと望んでいる。願わくば本稿が相互接続に参加してもよいと考えてくださる諸氏の参考になり、反応が頂けることを期待したい。

## 2 認知理論研究プロジェクト

### 2.1 なにを説明するモデルを求めているのか

これまで、遠隔地間でのコミュニケーションでは意思の疎通が充分に行なえないなどの問題点が指摘されており、その意味から映像や音声を用いたマルチメディアコミュニケーションシステム(動画音声情報通信システム、以下 AVICS と呼ぶ)には大きな期待が寄せられている。

しかしながら、メディア介在コミュニケーションでの問題点には単なる映像や音声の伝達では解決できないと考えられるものも少なくない。ネットワークを介して行なうコミュニケーションに比べて、直接対面して行なうコミュニケーションの方が柔軟性に富んでおり、有意義で集中的かつ効率的な会合を持たたとの印象を受け易いことを報告した研究はすでに散見される [1], [2]。我々の経験でも、直接対面して行うコミュニケーションではコミュニケーション自体や結果に対して満足感や安心感あるいは信頼感<sup>1</sup>を得ているように感じる。

こうした、メディアを介在したコミュニケーション

<sup>1</sup>ここでの信頼感は、工学の信頼性ではなく、共同主観性 [3] の要件である「承認」に近い意味である。

ンで感じる違和感や独特の感想(本論文では、仮にメディア不協和感と呼ぶことにする)を定量的に評価し、メディア介在コミュニケーションにおける問題を予測するためには、人間同士のコミュニケーションのモデルが必要になる。このようなモデルがなければ、たとえ十分な容量を持つ通信回線を利用したコミュニケーションメディアが登場したとしても、「結局、直接会った方が話が早い」「(メディア介在授業よりも)最終的には教室で教えなければならない」といった問題が未解決のままいつまでも残る可能性がある。適切なモデルを用いることで問題点を予想あるいは説明することが可能になり、その定量的評価と解決策の示唆を行なうことができるのである。

### 2.2 仮説

認知理論研究プロジェクトでは上述の問題点を説明するモデル構築が大目標となるが、モデル構築とその検証(実証的研究)は「にわとりと卵」の関係にあるので、仮説の生成とその検証を繰り返すことになる。我々はいくつかの仮説を準備しているが、ここでは次の二つについて述べることにする。

- 受信者の選択自由度とメディア不協和感の相関仮説
- 会話開始時の弱相互作用仮説

#### 2.2.1 受信者の選択自由度とメディア不協和感の相関仮説

従来のコミュニケーションメディアを越えるためのキーワードとして我々がもっとも注目しているのは受信者の能動性である。従来のコミュニケーションメディアでは受信者側の自由度が著しく制限されているからである。

古典的なコミュニケーションモデルであるシャノン=ウィーバーモデル(図.1)では<sup>2</sup>、データは発信側からの一方向的な流れなので、受信側が見たいデー

<sup>2</sup>あまりに古典的ではあるが、現在のほとんどのモデルはシャノン=ウィーバーモデルの拡張といっても差し支えない [4]。

タを選ぶことは表現されない(情報を完全に拒否することは可能であるけれども)。

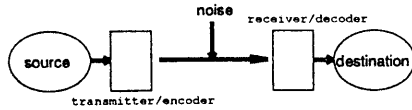


図 1: シヤノン=ウィーバーモデル

通常のメディア介在コミュニケーションではどの情報をどの経路で伝達するかは、基本的に非受信者が規定している。一般のテレビ番組では、カメラマンの映した映像、複数のカメラマンがいる場合にはディレクターの選んだ映像だけが放送される。受信者は自分の見たい映像を選択することはできない(もちろん、その番組を見ないことも可能であるけれども)。たとえ複数台のカメラのうちの任意のカメラを選びたくても、あるいは、特定のカメラの角度を上下左右に動かしたいと願っても、現時点では実現不可能である。

同様に、複数人(複数局)の参加するテレビ会議システムでも受信者には選択の自由はほとんどない。通信回線容量や利用者の画面面積の制約から、全員の映像を伝送することはできない。そのため多くのシステムでは、声の入感があった場所の映像をブロードキャストする(a)音声認識方式、議長の選んだ映像をブロードキャストする(b)議長方式、あるいは画面を小さく分割して全員を表示する(c)画面分割方式によって動画を伝送する。いずれの方式も受信者側に選択の自由はなく、受信者は非受信者に定められた映像を見ることしか許されない。

既存のコミュニケーションメディアは、ちょうど「覗き穴」からだけしか相手を見ることが許されていない状況である。日常の非メディア介在コミュニケーション(補足参照)での受信者が、一方的に受けとるだけの存在ではなく無限の選択可能性世界内を自由に移動する能動的な存在であることを考えれば、覗き穴でのコミュニケーションはメディア不協和感

の一因と推測される。大きなコストをかけて通信回線を大きくするより、通信システムの設計を工夫することで回避することが望まれる。

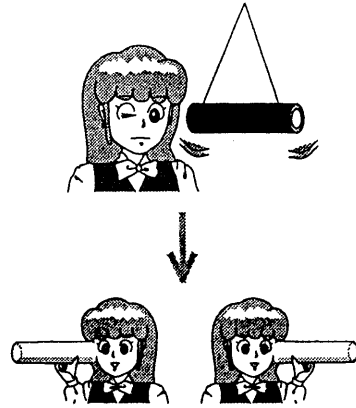


図 2: 覗き穴を意識せずに済ませるには?

ここで重要な点は「選択の可能性が制限されていないと受信者が考えることができる」ことである(図2)。

この仮説を検証するためには、受信者の選択自由度を高めたシステムを構築し、自由度を制限したシステムと比較すればよい。具体的には次の二つの技術を実装し、検討する計画である。

**話者自由選択** ある一定の(本研究では一人分の)動画しか伝送できないような細い通信回線でも、遠隔会議システムの全参加者を自由にみることができるようになる。どんなに大規模な討論会であっても同時にしゃべっているのは高々数人であろう。また、自分が注目できる人は高々一人であろう。したがって受信者が見たい画像(あるいは見ている画像)だけがつねに動画として転送されるようにすれば(あたかも、すべてが動画で送られているかのような)効果が期待できる(図3)。

受信者は他の参加者の動向に関わらず、自らの望む映像を得ることができる。もちろん一つの回線を共有する参加者が異なる送信元の映像を要求した場合には動画パフォーマンスは低下す

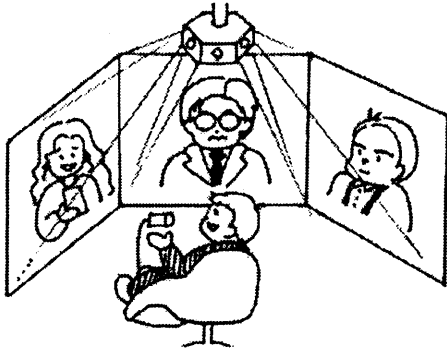


図 3: 見ている部分だけ常に動画…

るが、通常の討論ではばらばらの相手を見ることは少ないと予想されるので、実質上も問題はないと期待している。

システムの基本的なアルゴリズムは以下の通り。

(1) 送信者は最近接のルーターデーモンに対して送信する。(あるいは送信しない。) (2) 受信者は、最近接のルーターデーモンに(任意の送信元からの)データを要求する。(3) ルーターデーモンは送出元への経路途上のルーターデーモンに対してリクエストを出す。(4) ルーターデーモンは個々の受信者と受信者に至るまでの通信回線容量に応じてパーシャルサービス(映像の間引き)を実施する。

技術的な課題としては、(a) ルーターデーモンがネットワークポロジをいかにして知るか、(b) 経路をいかに管理するか、(c) パーシャルサービスのアルゴリズムなどが挙げられる。

**視点自由選択** オムニビューシステムは、セミ魚眼レンズによって撮影した映像の任意の領域を正規化して表示する機能を持っている [5][6]。図 4 は、オムニビューの機能を説明する図である。セミ魚眼レンズの入力映像は、垂直方向では a, b, c のように歪み、水平方向では d, e, f のように歪むが、オムニビューシステムを利用することにより、例えば P 領域を選択した場合にはリアルタイムで (B) のように正規化することが

できる。これを導入した AVICS は、(固定カメラであるにもかかわらず) あたかもカメラを移動させているように視線を移動することができる。通信回線がセミ魚眼レンズの情報を伝送するのに充分であるときは、各受信者が視線方向を自由に選択することが可能になる。受信者の視線方向の変更要求をいかに吸い上げるかが通信回線容量以外の技術的な課題である。

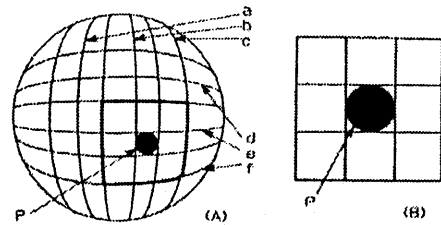


図 4: オムニビューシステムの機能

なお、従来のメディア介在コミュニケーションを「覗き穴」と表現したが、この特性が常に劣っていると主張しているわけではない。むしろ特定の場所を見せたいと意図する側にとっては役に立つし、ある特定のポイントに議論や注意を集中させたいときには有効な手段であることも予想される。重要なことは、メディアの特性を定量的に明らかにして使い分けの指針を得ることである。

### 2.2.2 会話開始時の弱相互作用仮説

どのようなメディア介在コミュニケーションでも、一方的に送信者が会話を開始しようとする。送信者には受信者の都合はわからない。

ところで、他者との交流が新しいアイデアを得るために重要であることは経験的に知られている。そのため、いくつかの研究所などでは個室を設けずにパーティションによる間仕切りを設けたことがある(パーティション方式)。

パーティション方式の良い点は、相手の都合を推し量ることができる点である。相手の忙しさを推察し、自分の用件の緊急度を検討することができるため、結果としてレアな状態のアイデアを相手にぶつ

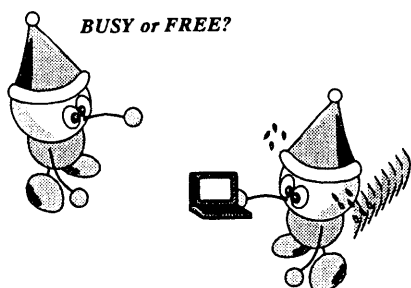


図 5: 会話を始める前に、弱い相互作用？

ける機会が多くなる(図5)。こうした、新しいアイデアを得るために重要な(けれども相手にとってはおそらく重要ではないだろう)インタラクションを電話で行うことは難しい。

そこで本仮説では、非メディア介在コミュニケーションでは情報交換のチャンネルが開かれるまでの微小時間内にも相手とのインタラクション(弱相互作用)があるのではないかと考える。

この仮説を検証するためには、弱相互作用を実現したシステムを実装し、アクセス頻度を評価すればよい。弱相互作用の実現方法はたくさん考えられるが、たとえば受信者空間への割り込みをいくつかのレベルで分けて用意しておく、送信者がそれを選ぶことができるようにしておく方法や、わざと画質を落とした映像からコミュニケーションを開始させる方法などが予定されている。

### 3 マルチメディア環境整備プロジェクト

本プロジェクトで実装する AVICS は、前節で述べたような仮説を検証するためのものであるが、いわゆる実験専用のシステムではない。本マルチメディア環境整備プロジェクトは、仮説の検証などには関係ないメンバーが日々のコミュニケーションを行う環境を提供することを目的としている。システムが実用的に利用されてこそ、実験室内実験では得られないデータを得ることができるからである。

## 3.1 なぜハードウェアから作成するのか

### 3.1.1 制御の自由度

研究者向けであることの最大の要件は「必要な仕様が公開されており、ハードウェアの制御の自由度が高いこと」である。ヒューマンインタフェース、バーチャルリアリティー、遠隔操作インタフェースなどに興味を持つ研究者は、要素技術として動画通信を自分のシステムに組み込みたい場合がある。1980年代後半から始まった CSCW 研究でも、動画通信システムを組み込んだ多くの協調作業支援システムが提案されてきた。しかし1990年代後半に入った現在でも、動画通信システムを自分のシステムに組み込むことは決して容易ではない。(ほとんどを海外からの輸入に頼っている)市販のハードウェアでは技術情報が十分に公開されないために制御の自由度が極めて低く、メーカーから提供された API からしか制御できないことが多い。したがって、研究として興味を引くような特殊な利用形態に対応することは実質上困難である。

もちろん、ボード自身が研究用途に対応できる仕様でなければならないことは言うまでもない。(しかし、市販品ではほとんど見あたらない。)

### 3.1.2 既存システムとの融合

同様の困難は、音声認識や画像認識システムの設計者にも生じている。ユーザ間の会話をモニタして適宜資料などを提示してくれたり、本人が不在の間でも簡単な情報交換をしておいてくれたり、あるいは自動議事録作成といった、秘書的作業を行なうシステムを研究するためには、AVICS を自システムに組み込んだり、対話環境でのリアルタイムデータがシステムに取り込めなければならない。しかし、個々のシステム自体に大変な労力を要するのにも、AVICS を組み込んでデータの相互交換を実現することはさらに大きなコストを必要とする。

このようなコストを低減させるためには、たとえば圧縮アルゴリズムを単に高効率であるからという理由だけで選択することはできない。音声のサンプリング周波数も既存の音声認識アルゴリズムがその

まま使えるという観点から選択されるべきであろう。

### 3.1.3 フリーソフト

一方、無償公開のソフトウェアはソースコードが添付されているものもあり、その意味から技術情報はほぼ公開されているとみなせる。しかし、実用的な性能を有していないものがほとんどである。それはソフトウェアだけで AVICS を実現することの限界でもある。ハードウェアから準備された AVICS が要素技術の一つとして提供されれば、インタフェース関連研究者にとっては大変有意義なことといえる。

## 3.2 ハードウェア

図6は、我々が計画しているボードと、その構成を表している。

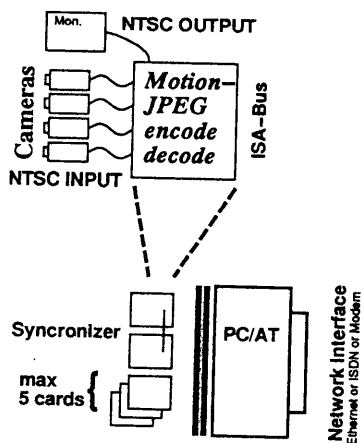


図6: JPEG 圧縮伸張ボード及び構成

- PC の ISA バスに接続して利用する。
- NTSC 入力を 4 つ、NTSC 出力を 1 つもつ。入力ポートはソフトウェアで切り替える。
- NTSC 入力画像を Motion-JPEG で圧縮することができる。
- ソフトウェアによって圧縮と伸張を切り替えることができる。

- ハードウェアレベルで同期機構を備え、ボード二枚を利用してステレオ送受信ができる。

- 一秒あたりの枚数、画像一枚ごとの圧縮率をソフトウェアで制御できる。

これとは別に音声・通信ボードも計画している。

- PC の ISA バスに接続して利用する。
- マイク入力は 4 つ、スピーカ出力を 1 つもつ。
- 入力は 16kHz サンプリング、ADPCM。
- 圧縮伸張を並列に行う。
- 同時刻にサンプリングされた別個の場所からの音声データを合成する。
- ISDN ポートを持つ。

市販のボードでは NTSC 出力を持たないものが多いが、NTSC への出力は研究上重要である。たとえば等身大に相手を表示することによって心理的距離感が近くなることが報告されているし [7]、ヘッドマウントディスプレイを利用したいときには必要となる。

画像圧縮技術として将来は MPEG が利用されるであろうと考えられるが、加工が容易であること、他のシステムとの融合が容易であること、設計コストを押さえられること、Ether などのアーキテクチャでは MPEG よりも JPEG の方が安定して伝送できると期待されることなどから、Motion-JPEG を採用した。

計画通りに作業がすすめば、これらのボードは 9 月に利用可能となる。

### 3.3 OS とリアルタイム性

システム全体が安価になるよう PC で利用し、さらに多くのユーザが利用できるように OS は DOS/Windows/Unix、ネットワークには Ethernet/ISDN/modem を予定している。

すでにマルチメディア通信を念頭においたリアルタイム OS の研究は多く見られ、そのような最先端 OS 技術を投入すれば一秒あたりの画像枚数は若干

増えると期待されるものの、(a)現状の通信回線では(容量が低過ぎるためにOSでの努力の)影響が少ないと予想されること、(b)たくさんのユーザがいることこそ重要であることを考慮し、我々は厳密なリソース管理を追求しない戦略を採用した。例えば、我々の採用したISAバスでは完全なリアルタイム制御が困難である。しかしEthernetやISDN、公衆回線といった貧弱な回線での利用ではリアルタイム性の追求には限界がある。また、音声と映像の同期は考慮していないが、これも映像のクオリティの追求には限界がある現在の通信回線では(同期の問題は)致命的ではないと予想している。

### 3.4 研究課題

以上のシステムでは、認知科学研究以外の研究テーマを扱うことができる。その例としては以下のような項目が計画されている。

**ロボット遠隔操作インタフェース** 物理的運動を伴う作業では一般に視覚フィードバックを利用しており、奥行き感覚を得るためにはステレオ画像の伝送とヘッドマウントディスプレイによる立体視が有効である(図7)。ここではインターネットを介してステレオ画像を伝送する上での課題を検討する。

**議事録作成検索支援システム** テレ討論会の実施によって、音声動画データは爆発的に増大する。データの容易な再利用が電子メディアの最大の利点であることを考えれば、これらの膨大なデータを如何に整理し、容易に検索できるようにするかが重要なテーマとなる。

整理検索に求められる要件は(a)人間の記憶想起メカニズムの理論的裏付けに基づくシステムであること、(b)データの可搬性(および加工性)に優れた表現形式であることの二点である。

**リソース管理技術** 本プロジェクトで準備するハードウェアは厳密な意味でのリアルタイム性を追求するには不十分な面もあるが、リアルタイムOS上のサービスを研究するためのテストベッドとしては充分に利用できる。例えばフレーム単位

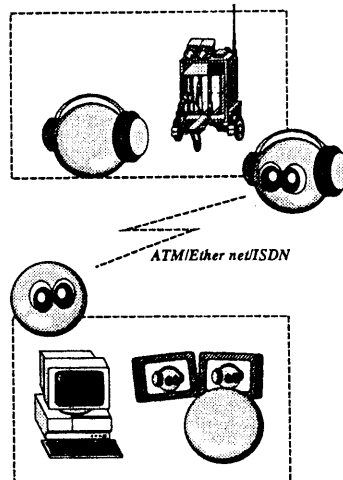


図7: ステレオでの送受信

での制御が可能のため、ネットワークの利用状況に応じた圧縮を行なうことができる。

その一方、音声と画像の同期を考慮しないことで、実用性が高かつ柔軟性に富むパーシャルサービスも期待される。また、アプリケーションを遠隔討論に限ることで効率の良いパーシャルサービスが期待される。例えばコミュニケーションでの利用では相手の顔の部分が重要になるため、全体の解像度を320x240から80x60などに落してデータを間引くのではなく、解像度を落さずに顔領域の80x60を切り出して転送することが考えられる。

この他、JPEGの圧縮率は画像によって異なるという性質を利用して、圧縮率に変化があったときには画像自体になにか大きな変化が生じたと推測することができる。このようなデータを優先的に配送したほうがよい可能性がある。

## 4 まとめ

ロジャースは過去のコミュニケーション研究の方法論にはいくつかの欠点があると指摘している[8]。その一つは「新通信技術の設計開発過程で、最大の

貢献を果たすはずのコミュニケーション研究者が、実際に参画するのは極めて遅い時点(設計が終了してしまっただけ)である」というものである。事実、コミュニケーションのこれまでの研究は、既に流布しているメディアを対象にすることがほとんどであった。本稿で紹介したプロジェクトは、そのような批判への再反論として、新しいコミュニケーション研究の流れをつくり出す試みでもある。

今後、非メディア介在コミュニケーションとメディア介在コミュニケーションとの効果の差異について、特に遠隔教育システムや医療相談や精神分析、心理分析を含む遠隔コンサルテーションシステムなどに興味を持つインタフェース研究者によって定量的な議論がなされることを期待したい。

## 謝辞

本プロジェクトの遂行にはさまざまな方々の御協力を頂いております。特に電子技術総合研究所情報科学部 諏訪 基 部長のご支援に感謝致します。また、NTT マルチメディア共同利用実験プロジェクトにおいて共同研究を予定している慶應義塾大学安西祐一郎先生および研究室のスタッフの皆様、新潟国際情報大学松井孝雄先生、ハードウェア試作にご協力頂いている株式会社オムニビューシステムの皆様にも感謝いたします。

## 参考文献

- [1] Fowler, Gene D. and Wackerbarth, Marilyne E. "Audio Teleconferencing versus Face-to-Face Conferencing: A Synthesis of the Literature", *The Western Journal of Speech Communication* 44, pp.236-252, 1980.
- [2] これからの社会における電気通信の役割, 財団法人電気通信総合研究所報告書, 1976.
- [3] 廣松 渉, 増山真緒子, "共同主観性の現象学", 世界書院, 1986.
- [4] 森岡正博, 意識通信, 筑摩書房, 1993.
- [5] マルチ映像システム インタラクティブ・メディアへの応用, 通信と放送 融合へのシナリオ (日経ニューメディア別冊), 日経 BP, pp.160-169, 1994.
- [6] 山地正城, パラダイムの転換をもたらす新デジタル映像システム, 映像情報, Vol. 26, pp.29-35, 1994.
- [7] Gi seok Jeong, 市川裕介, 田中俊介, 岡田謙一, 松下温, 多地点会議システム MAJIC の評価に基づく新システムの提案, グループウェア研究会報告 8-2, pp.7-12, 1994.
- [8] E.M. ロジャース (安田寿明 訳), コミュニケーションの科学, 共立出版, 1992.

## 補足

本稿では、メディア介在コミュニケーションと非メディア介在コミュニケーションの対比を検討したが、これらの言葉「メディア介在コミュニケーション」と「非メディア介在コミュニケーション」が適切に定義できるかどうかは、極めて重大な問題である。なぜならば、この言葉の定義が(本稿で用いられたように)「メディアを介在していればメディア介在コミュニケーションであり、そうではなくて直接対峙しているときのコミュニケーションは非メディア介在コミュニケーションである」というものであれば、この両者は(定義により)永遠に相容れないものであり、メディアを介在したコミュニケーションにおいてメディアを介在しないコミュニケーションと同等の効果を得る、という工学的目標も(定義により)実現不可能になるからである。

また、日本語でメディアといえば何らかの電子通信技術が介在しているものを指すことが一般的だが、たとえば音声による対面会話であっても空気という media を媒介したコミュニケーションと捉えることができる。

その意味から、本稿で用いた用語は議論の立脚点としては適切ではないかもしれないことを付記しておく。