

# 記憶プロセスに基づいた対話プログラム用の データベースのダイナミック管理システム

A dynamic memory management system based on forgetting and recalling.

ミハウ・プタシンスキ

Michał PTASZYŃSKI

g19003@edu.otaru-uc.ac.jp

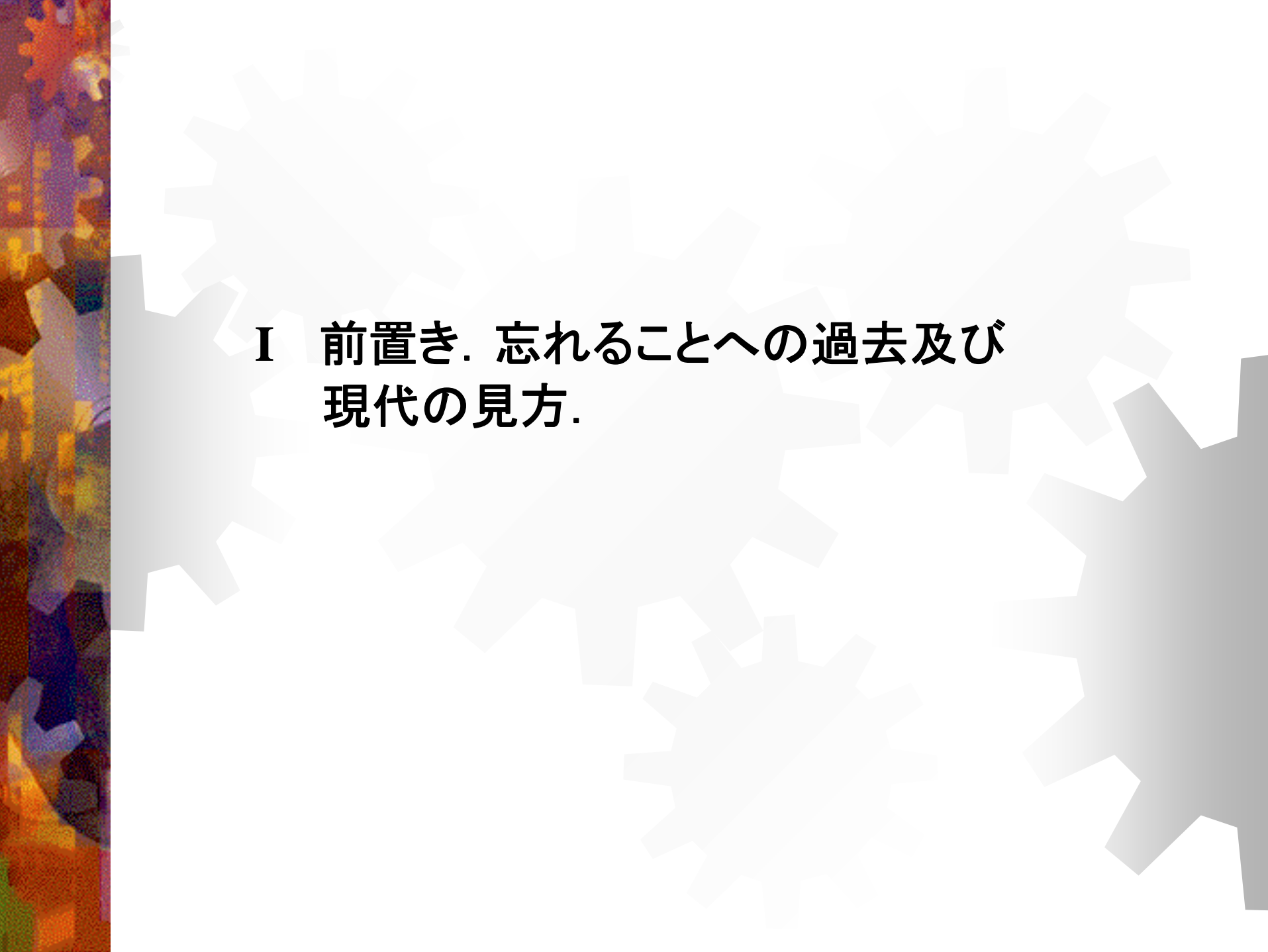
指導教官: 佐山公一 准教授

## 概要

対話プログラムにおける会話の不自然, 不適切さなどのところを直す方法を考え, 人間の記憶に基づくデータベース管理システムが提案される.

本システムでは, 人間の忘却 (forgetting) と回顧 (recalling) の特性をデータベースのダイナミックな管理に利用し, 機械的に考えた.

機械による忘却・思い出すアルゴリズムは, コンテキスト・ユニットの使用頻度及び感情的分析から求められた感情の値によって決定される.

The background features a light gray field with several faint, overlapping gear shapes. On the left side, there is a vertical strip with a colorful, abstract, and textured pattern in shades of orange, red, purple, and blue.

# I 前置き. 忘れることへの過去及び現代の見方.

① 忘れることについての固定観念:

・忘れることは、人間の頭の中の欠点、短所。

・物忘れしたる、苛苛したりす。

・忘れることさえなければ、人間は何でも覚え  
つづけることができ、日常生活が楽だった。

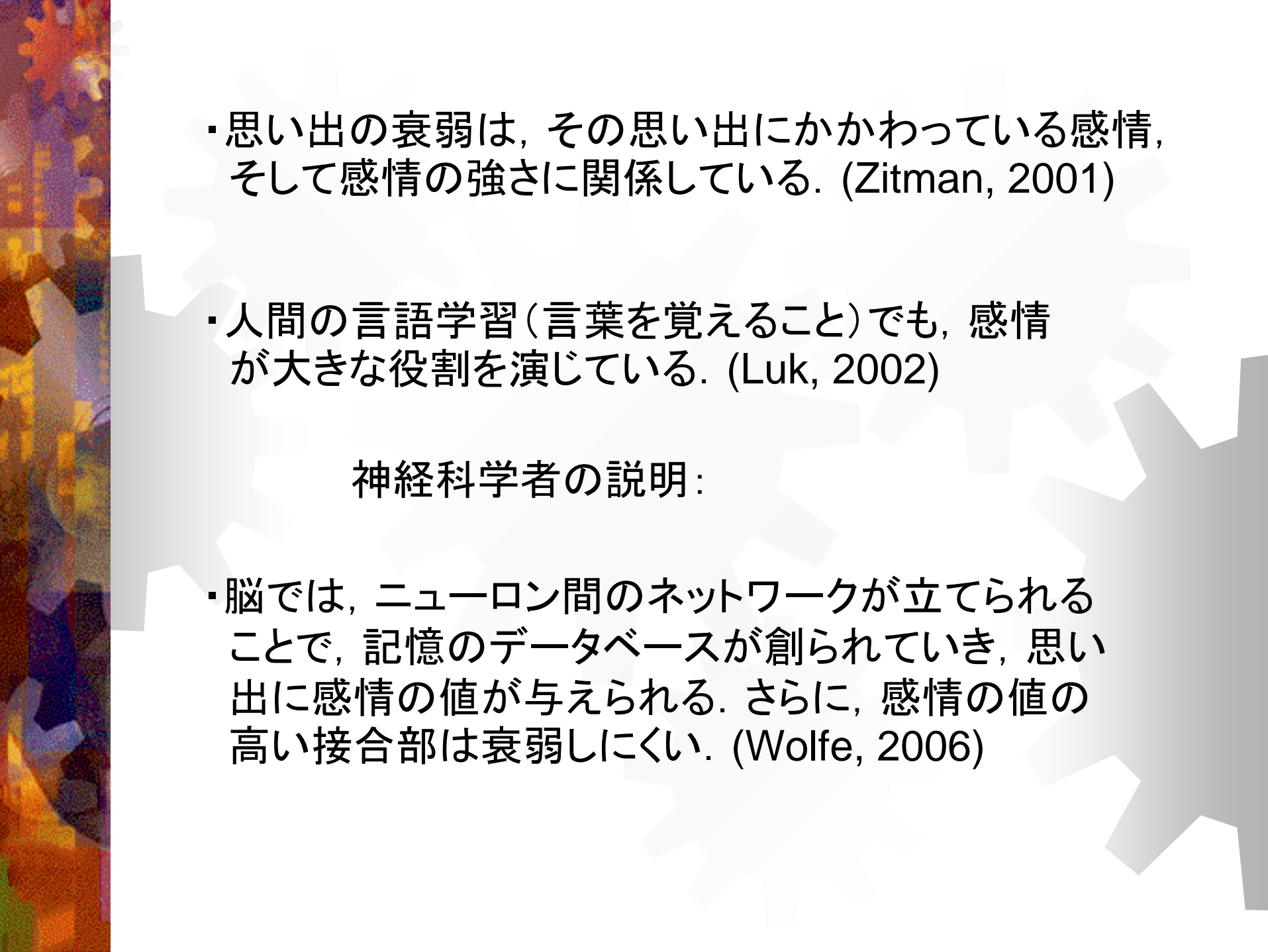
忘れること一悪心

## ② 忘却についての従来の知識:

- ・忘れることは、絶えず続くプロセスで、年代順に進んでいく。(Ebbinghaus, 1885)
- ・忘れるプロセスでは、知識が再整理・再配列され、その部分が不活性になり、部分へのアクセスが遮断される。(Anderson, 1983)
- ・その知識は、認知科学の上に、工学や人工知能の研究にも利用されたことはあった。(Markovitch and Scott, 1988)

さらに、神経科学者の説明:

- ・脳内には、2つのニューロン(神経細胞)を新しく接合したシナプス(接合部)は、接合直後から衰弱していく。(Waddell, 2004)

- 
- ・思い出の衰弱は、その思い出にかかわっている感情、そして感情の強さに関係している. (Zitman, 2001)
  - ・人間の言語学習(言葉を覚えること)でも、感情が大きな役割を演じている. (Luk, 2002)

### 神経科学者の説明:

- ・脳では、ニューロン間のネットワークが立てられることで、記憶のデータベースが創られていき、思い出に感情の値が与えられる。さらに、感情の値の高い接合部は衰弱しにくい. (Wolfe, 2006)



### ③ 忘却についての最近の発見:

忘れることは、損ではなく、知識取得を促進させる有益なプロセスだと、前から論理的に仮定していた研究者はいたが・・・ (Markovitch and Scott, 1988)

### 最新発見:

忘却は欠点ではない!!

逆に、知識と記憶を整理、配列させることで、大切なことを覚えるために貢献する. (Kahn et al., 2007)

忘れること ≠ 悪

## 忘却への新しい見方：

忘れるというプロセスは，消化，呼吸などと同じく，自然なプロセスである。

だとするなら……

忘却プロセスを機械的に算定できる。さらに，

その特性を利用し，機械でもプラスに働くように効果的に導入できる。





## II アイディアの誕生.



## ① 対話プログラムのスピーチの不自然性問題

対話プログラムにおける研究がますます進んでいく一方、スピーチの不自然さ、不適切さという問題はまだまだ乗り越えていない。

その原因は、自然な会話の文は、文脈に依存するが、コンピュータで処理を行なうと、文脈自由の単文の処理は可能だが、文脈で結ばれた多数文を考慮しようとする、コンピュータで処理することが困難である。(荒木, 2004)

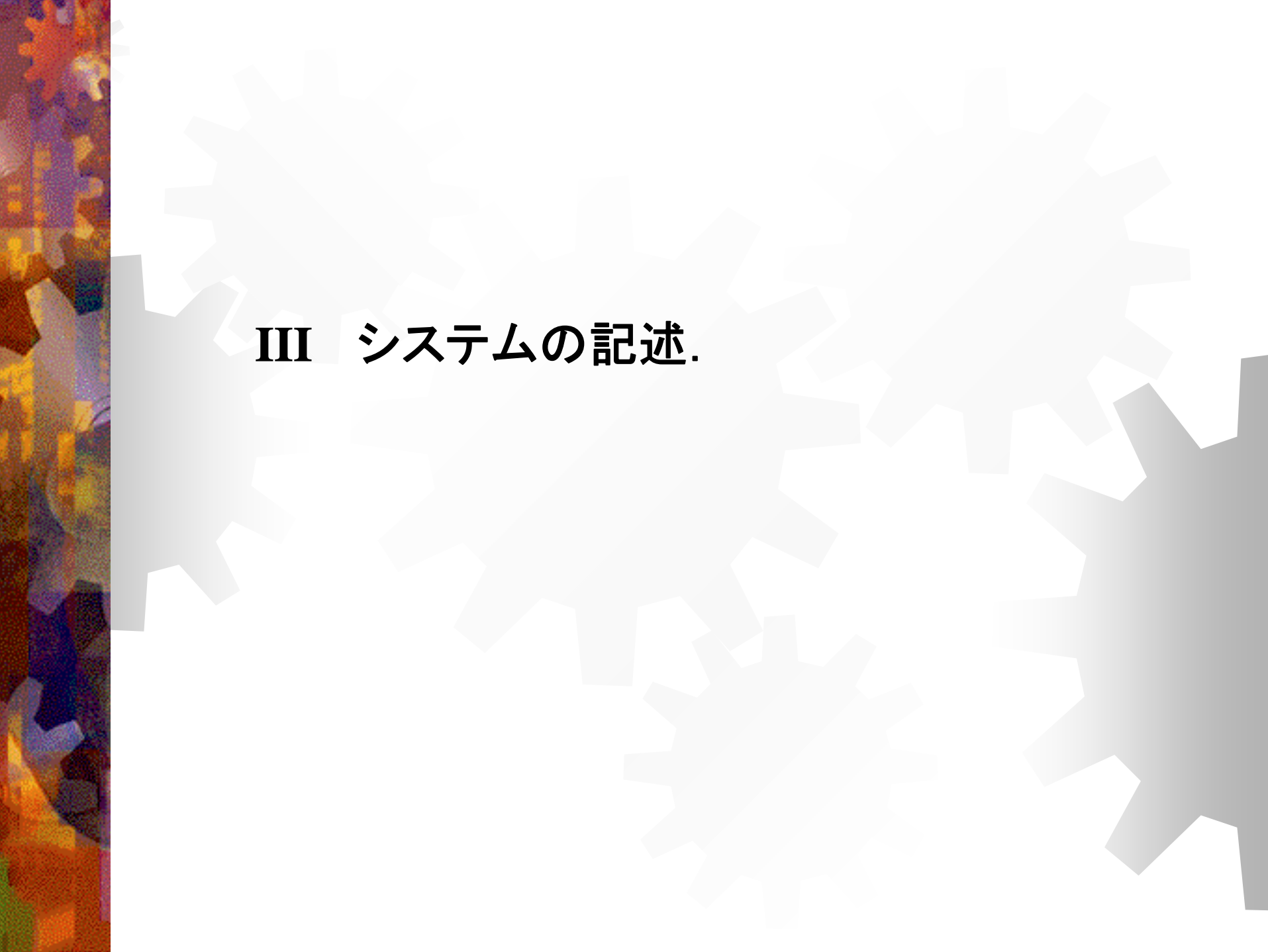
## ② 自然言語処理—より自然な言語の処理を目標に！

機械のスピーチの不自然さ問題の解決方法を考え、会話プログラムの“記憶”を人間の記憶を真似して管理するシステムのアイデアを提案したい。

このシステムは、人間の記憶メカニズムの要素、すなわち、忘れること、そして思い出すことに基づくアルゴリズムを利用し、プログラムの“記憶”つまり、会話のデータベースをダイナミックに管理する。

忘却アルゴリズムは、忘れることの代表的な特性に基づくと計画されている。それらは、下記のようなものである。

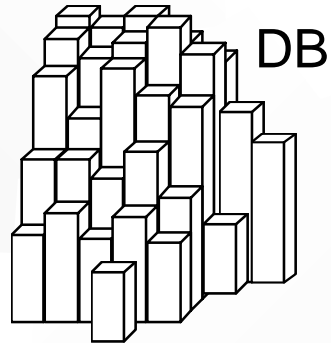
- ・記憶は拡大していくデータベースである。
- ・忘れることは、記憶を整理、配列させることで、知識取得の必要な要素である。
- ・忘れるプロセスは絶えず続く。
- ・一つの思い出、つまりコンテキストの枠は、連想 (association) が設置した直後から忘却の候補者になる。
- ・忘却は、連想の使用頻度、そして
- ・会話の感情分析から得られた感情の値によって決定される。

The background features a light gray field with several large, semi-transparent gear shapes of varying sizes and orientations. On the far left, there is a vertical strip with a colorful, abstract, and textured appearance, possibly representing a gear mechanism or a system component.

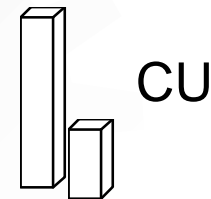
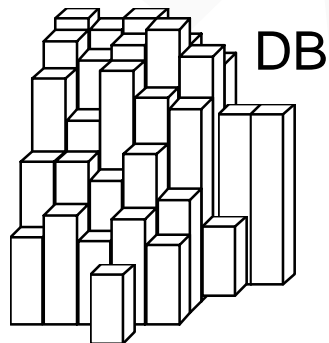
### III システムの記述.

# 1. 独立の会話プログラムの働き方の全般的記述.

- ・データベース(記憶の同等物).



- ・前の会話は, 一つ一つ, コンテキスト・ユニット(CU)としてデータベース(DB)に記録(インターネット・メッセンジャーのログのように).





- ・ユーザが話をかけたら、新しい会話が始まり、さらに、同時に更新のコンテキスト・ユニットとしてDBに記録.

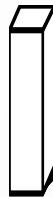
A: すまん。俺も裏ぐった。文才が無いから、過程は書けないけど。このスレまで魔力ありすぎ... おまいらにも光あれ...

B: なんだとこあんちくしょうあああああ

: きになる

B: >>731 彼女が出来たのか？

A: >>734 違うけど。でも大チャンス。こういうこと嫌くとネタにしか聞こえないよな。とにかくおまいら外に出てみる



- ・一つの会話記録 = 一つのCU.

A: すまん。俺も裏ぐった。文才が無いから、過程は書けないけど。このスレまで魔力ありすぎ... おまいらにも光あれ...

B: なんだとこあんちくしょうあああああ

A: きになる

B: >>731 彼女が出来たのか？

A: >>734 違うけど。でも大チャンス。こういうこと嫌くとネタにしか聞こえないよな。とにかくおまいら外に出てみる



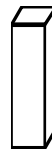
A: 帰還しますたとりあえず着替えてきます(・▽・)

B: 電車キタ——(▽)——!!!!

A: 今戻りました。

結果から報告すると成功でうえdrftgyふじこ  
思い出しながら経過など報告させていただきます。  
それとみんな有難う。

B: 性交キタキタキタ——(▽'≡(▽'≡'▽)≡'▽)——!!!!!!!



A: こんにちは

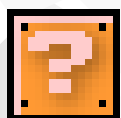
B: さよなら

USER:



# ・ユーザの発言をしたら・・・

こんにちは、今日、サッカーをしました。



[サッカー][助詞][N3]

プログラムは、適切なコンテキスト  
連想についてDBに質問する(query).  
コンテキスト連想={n-gram, 会話手続き等}

...

A:さつまいもってうまいな

B:ほんまだね。

A:彼女と一緒に夜空を突ながら食うのは最高だな！



...

...

A:なんてことだ。

B:信じられないやろ？

A:お巡りさんに殴られるなんて。

B:サツってそんなもんだよ。



...

...

A:いっぱい走ったよ。

B:おつかれ。

A:へとへとだよ。やっぱサッカーは疲れるな。

B:じゃ、生よりゲームやれば？



...

・DBへの質問の結果を元に発言する:

疲れましたか。



ただし・・・

・適切な連想を見つけられない場合もある。



そのときは、外のDB(例えばインターネット)に質問し、



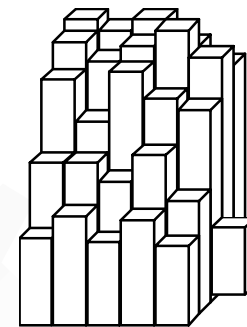
連想リストを取り出す。



・DB及びインターネットへの質問の結果は、CUとしてDBに記録される。

その区別は:

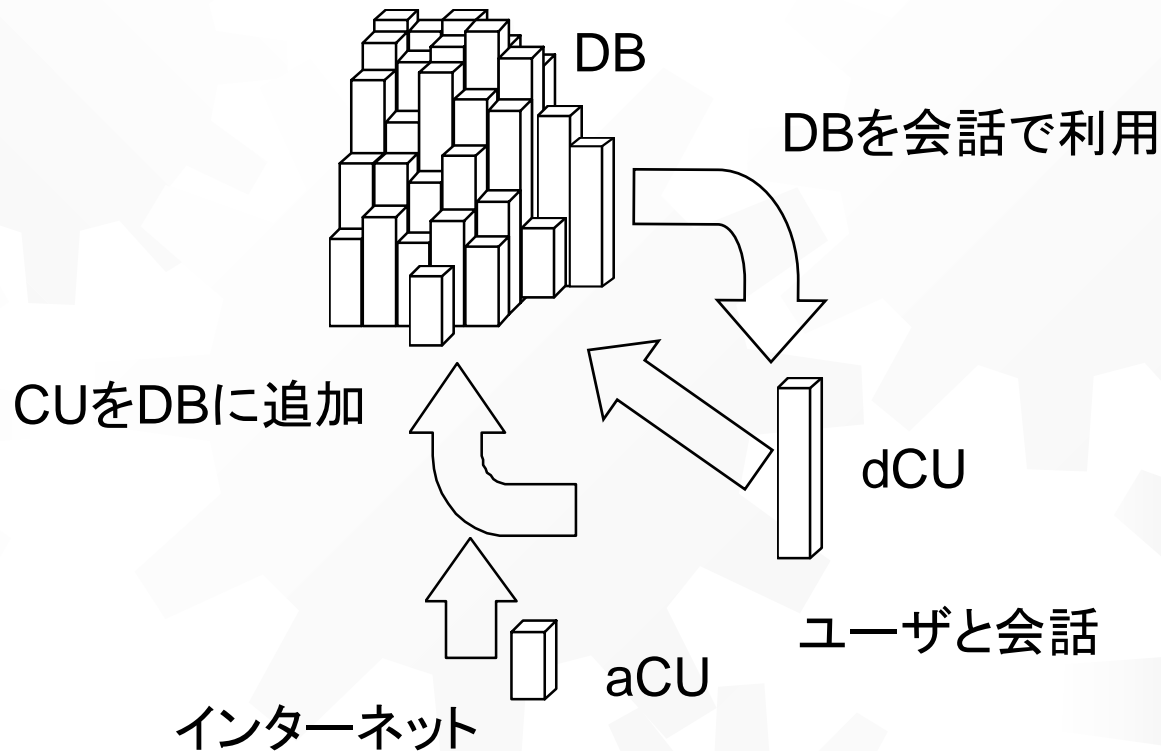
- ・会話ならなるCUは、単語、会話手続き、感情の値を含む。(dialogue CU - dCU)
- ・インターネット検索の結果からなるCUは、単語の連想リストのみを含む。(associative CU - aCU)



しかし、プログラムは、両方を同じように扱う。



## データベースの拡大化プロセス:



しかし、このようにデータベースは巨大に拡大し、処理しにくくなってしまふ。この問題を解決するためには、忘却アルゴリズムが導入される。

## 2. 忘却アルゴリズム.

前述のように、忘れるプロセスでは、記憶が再整理・再配列され、その部分が不活性になる。ここでは、データベースのそれぞれの部分(CU)へのアクセスが遮断されることを意味する。

アクセスを遮断するかどうかは、CUの。  
**使用頻度**及び**感情的分析**から求められた**感情の値**によって決定される。



## 使用頻度

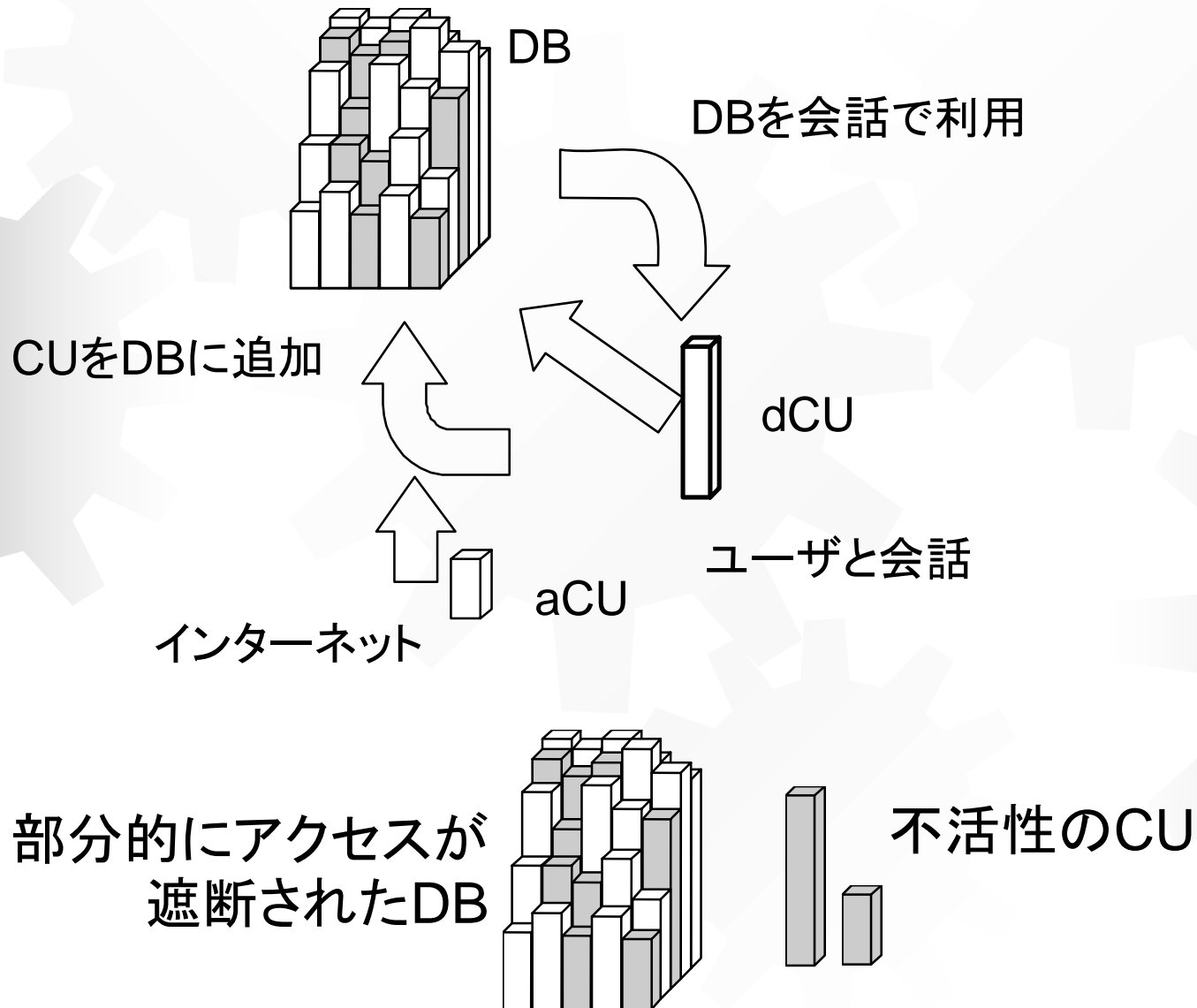
DBの中には、頻繁に利用される連想(CUの内容)があれば、めったに触れられていないものもあるため、忘却プロセスをそれに基づくことは難しくない。つまり、各連想が会話で利用されていなかった場合は、それらを含むCUは忘却される候補者になる。

## 感情の値

感情の分析と値については、本発表の次の部分で説明する。



# 忘却アルゴリズムの適用したシステムの図:



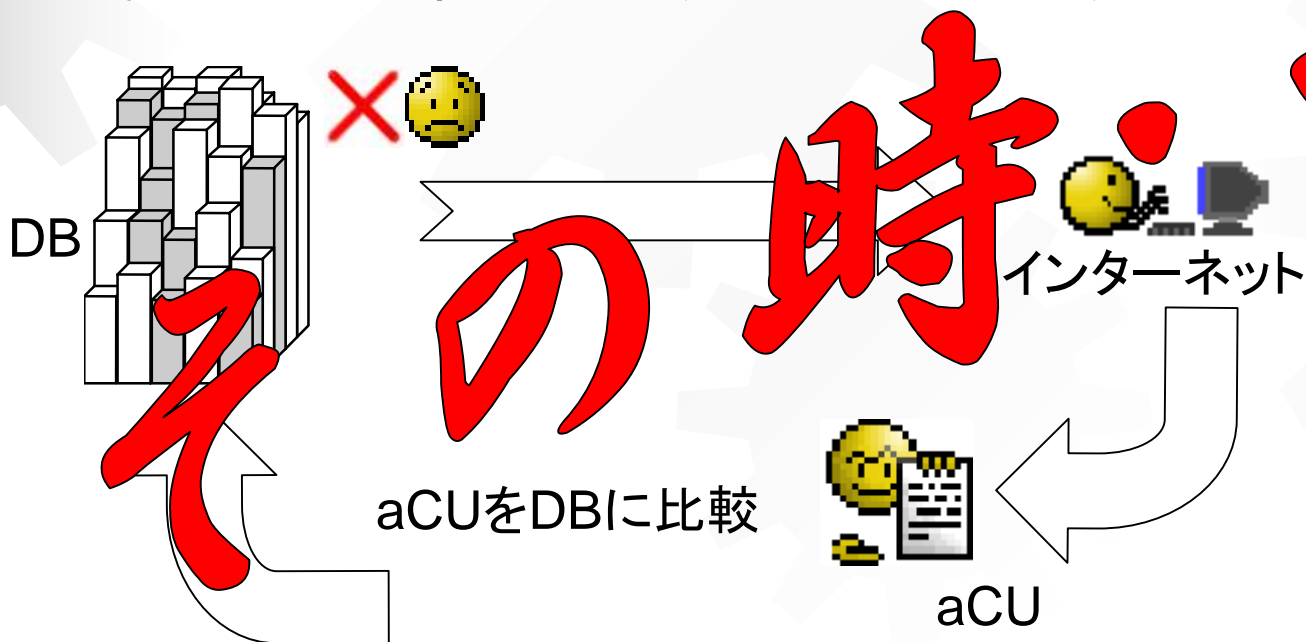
### 3. 回顧アルゴリズム.

以前の人工知能研究では、知識を忘れてプログラムの能力を上げようとするアイデアが内気に提案されたことはあるが、それらは、主に グラフ検索用の簡単なプログラムにおける知識を集める段階でのリンクの重みの削除 (Markovitch and Scott, 1988), または、理論的なコネクショニスト・ネットワークにおける、ヒューリスティックスの削除 (Ishikawa, 1990) を指す。 その意味で、実用に注目した単なるDB管理方法という本アイデアとは異なる。

ちなみに、以前の忘れるメカニズムを利用しようとしたアイデアでは、思い出の回顧、つまり、忘れた情報の思い出し、人間のにとって当然にあることが**無視**されていたようだ。

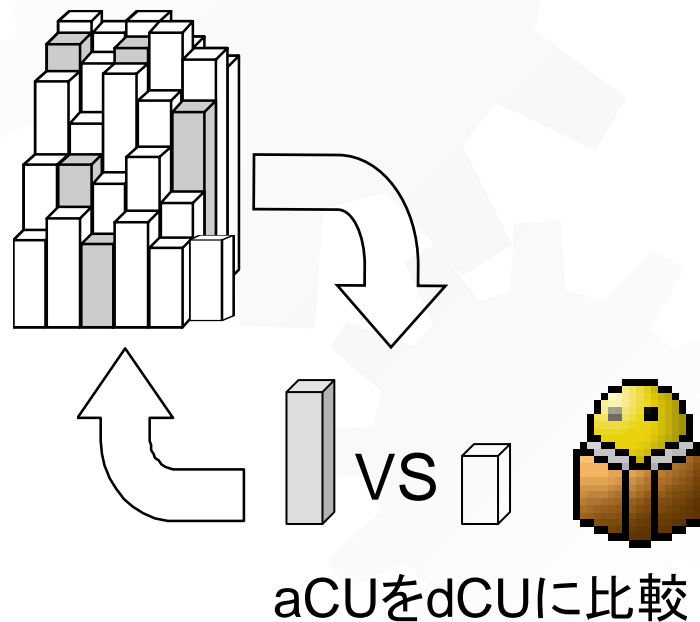
## 回顧アルゴリズムでの段階:

- ・利用されていないためアクセスの遮断したCUがある.
- ・そのため, DBにはあるが, “見えない” 連想がある.
- ・インターネットに質問.
- ・連想(単語)リスト獲得(aCU).
- ・そのリスト内容を忘却済みのCUと比較.



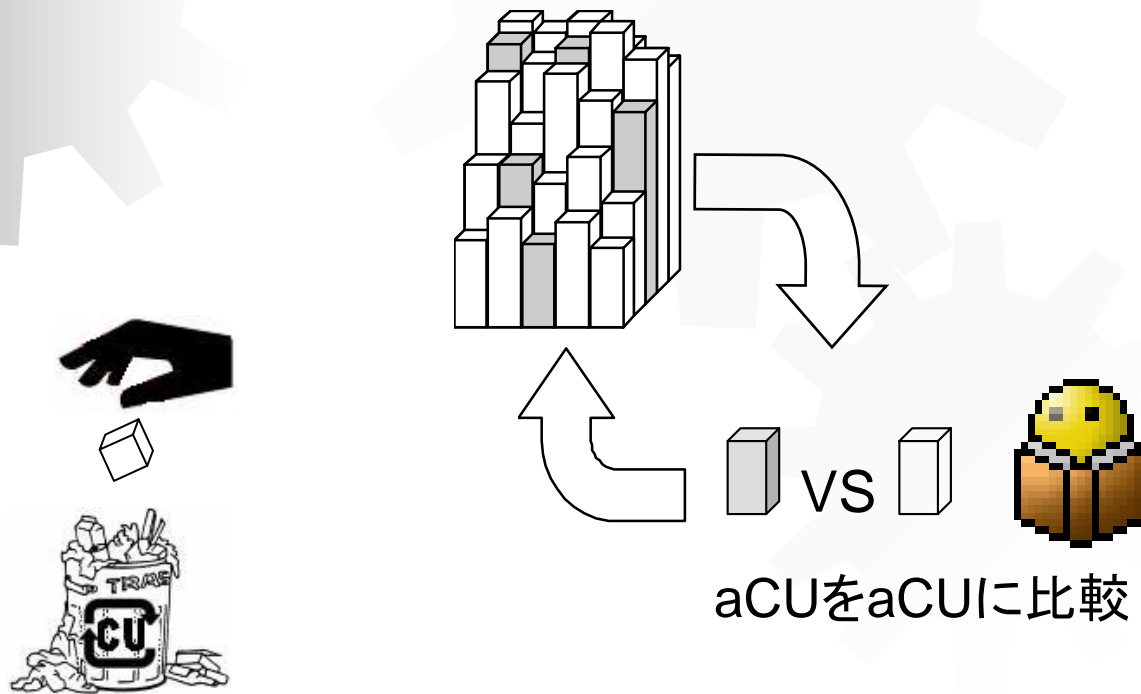
# インターネットから獲得したaCU⇒ DB内の忘却済みのdCU及びaCUと比較.

- ・忘却済みのdCUと連想する場合：  
そのdCUが回顧され，活性化され，アクセスが再開される。  
新しく獲得したaCUは，活性的aCUとしてDBに追加される。



・忘却のaCUと連想する場合：

二つのaCUの内容が比較され，必要な場合に新しい内容が追加される．重なる残りは削除する．内容が同じな場合は，新しく獲得したaCUがそのまま削除．比較されたaCUが回顧される．





#### 4. 感情の分析と値.

前述のように, 人間における忘却には, 思い出に与えられた感情が関係し, 感情の値の高い接合部は衰弱しにくい.

では, 感情の分析と値を与えるメカニズムは, どうやって本研究で適用するのでしょうか.

**実は, 簡単な, しかも 確実な方法がある!!**





本研究では, CUが思い出として考えられる.

ということで, 感情の値は, 各CUに与える必要がある.

具体的には, 感情の値 (Emotive Value - EV) がCUの中の各発言に与えられ, CUの全般的な感情の値は, その近似値になる.

さて, この値は, どこから求められるのか...

本研究では、対話プログラムとの会話が対象となる。  
このような会話は、ユーザがキーボードで書き込みを  
することによって行なわれる。

普段、会話では、言語以外に、身振り、しぐさ、姿  
勢、表情、イントネーションなどが利用される。

しかし、対象となる会話方法では、人間間コミュ  
ニケーションの本の一部しか伝達できない。

学者によっては、詳細な意見が異なるが、言語  
的要素のみで表現できるのは、一般伝達の10%  
を超えていないという。(Szopski, 2005)

この極限の環境では、感情を機械的に分析したい  
のであれば、魔法の業でも使う必要があるのでは  
ないか… …と想ったら、実は、

言語学の**基本的な知識**が充分である。

言語学では、文が感情に色彩されているかどうかは、下記の要素の有無によって確実に決定できる。

**間投詞** すげえ, うおお, やったな 等

**感嘆文** なんという~, これが~ 等

**擬態語** ワクワク, ドキドキ, (胸が)ぐじゃぐじゃ(する) 等

**愛称** よっちゃん(良子), 黒ちゃん(黒川など), まりりん(真理子)等

**くだけた言い方** しています → してんの 等

**卑語** 分からねえ, よく降りやがるな, おっぱじめるぜ 等

**感情の語彙目録** (名)喜ぶ, 悲しむ, むかつく  
(形)嬉しい, 悔しい 等

**音声的な要素** モーラの延長, 「!」, 「?」, それらの複数使用 等

具体的な例:

- ・これは本でしょ? (EV=0) → ・こりゃ本じゃろ!? (EV ≠ 0)
- ・彼女は、大きいかさをもってきて、信之介を強く殴った.  
(EV=0)  
↓
- ・あいつあ でっけーかさをもってきやがって、シンちゃん  
にひでーボコボコいれ ちまった. (EV ≠ 0)

このように、発言ごとに、感情の値が決定され、CU内の発言ごとに与えられる。

ちなみに、dCUとaCUが同じように扱われると述べたが、aCUは、会話の記録ではなく、連想リストに過ぎないため、ほとんどの場合「感情値＝0」という風に考えられる。

The background features a light gray field with several faint, overlapping gear shapes. On the left side, there is a vertical strip with a colorful, abstract, and textured pattern in shades of orange, red, purple, and blue.

## IV システムの導入計画.



システムは、まだ理論的な段階にあるが、その部分が既にできていると考えられよう。

現在存在している多数の対話プログラムのどれかを利用する可能性はある。

決定 →

予備試験段階では、本研究室の樋口の「対話システム」\*を利用できることが望ましい。



## 感情の分析:

プタシンスキ(2006)\*\*の感情表現の分析に基づき, 発言内オリジナル感情分析システムを製作, 提案したい.

## 感情の値:

初期の段階では, 感情色彩の有無のみに基づく(値は0か1). 後は(予備テスト段階後), 値をもっと詳細な方法を考える. 例えば統計学的に発言内の感情色彩表現の割合など.

\*\*) ミハウ・プタシンスキ(2006)『萌える言語. インターネット掲示板の上の日本語会話における感情表現の構造と記号論的機能の分析・「2ちゃんねる」電子掲示板を例として・』, ポズナン: アダム・ミツキエヴィッチ大学

## 予備テスト段階

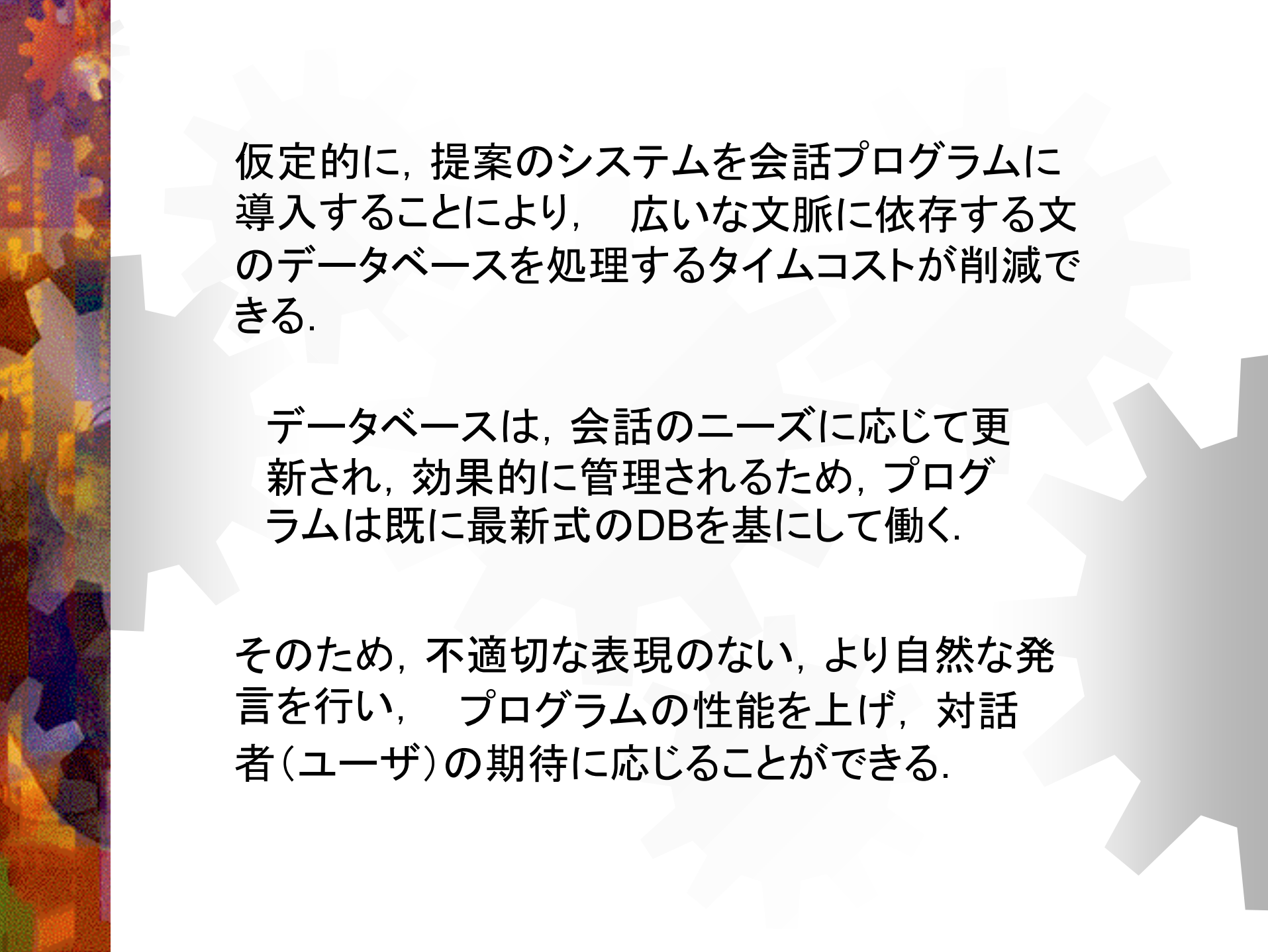
忘却・回顧アルゴリズムのオン・オフ→性能, 効率, 処理にかかる時間などを観察.

## 実行テスト段階

1. ユーザがプログラムと会話をし, それを評価する(チューリング・テストのスタンダードなアイデア)
2. プタシンスキ(2006)の会話分析への見方に基づき, テスト中の観察者の逆説 (Observer's Paradox) を省けるアイデアを利用する.  
すなわち, 対話プログラムが他の対話プログラムと会話をし, その会話の記録はユーザが評価する.

The background features a light gray field with several large, semi-transparent gear shapes of varying sizes and orientations. On the far left, there is a vertical strip with a colorful, abstract, and textured appearance, possibly representing a gear mechanism or a mechanical part.

## V 結論.



仮定的に、提案のシステムを会話プログラムに導入することにより、 広いな文脈に依存する文のデータベースを処理するタイムコストが削減できる。

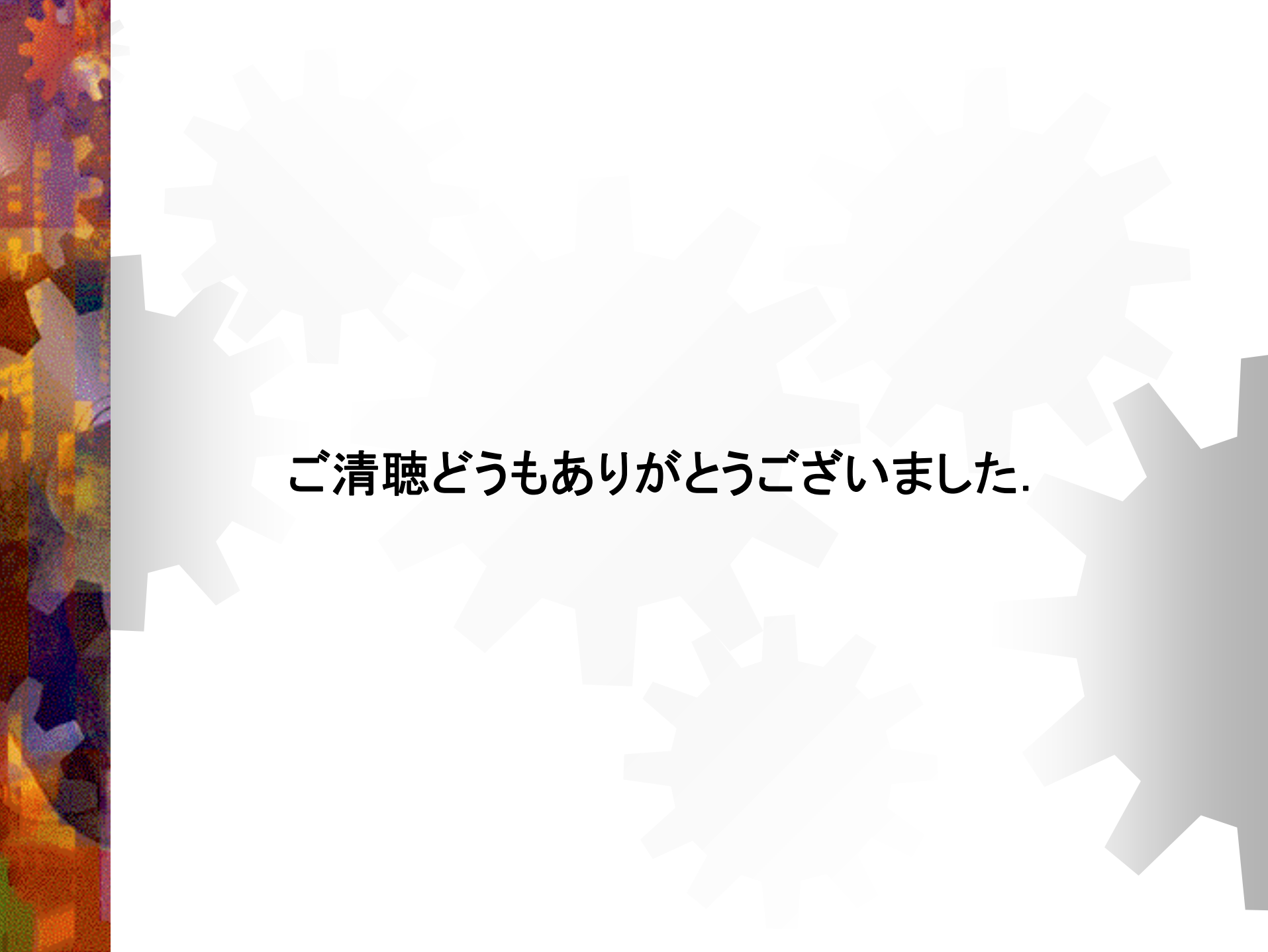
データベースは、会話のニーズに応じて更新され、効果的に管理されるため、プログラムは既に最新式のDBを基にして働く。

そのため、不適切な表現のない、より自然な発言を行い、プログラムの性能を上げ、対話者(ユーザ)の期待に応じることができる。

## 有効性の見込み

- ・人間との会話ができるプログラムを作成.
- ・広いコンテキスト処理の研究への貢献.
- ・人間の記憶のモデル化： 事項知能研究  
心理学  
認知科学 等へ貢献.
- ・プログラムへの実用：  
顧問プログラム  
ストレス解消用の雑談プログラム  
カウンセラープログラム 等.
- ・会話の中の感情分析によって：  
人間の感情の研究への貢献.

等...

The background features a light gray field with several faint, overlapping gear shapes. On the left side, there is a vertical strip with a colorful, abstract, and textured pattern in shades of orange, red, purple, and blue.

ご清聴どうもありがとうございました.





つづく