

イソップワールド - - - 対話処理に向けて

杉野昇史, 村上浩規, 目良和也, 徳久雅人, 岡田直之

九州工業大学情報工学部

820 飯塚市川津

あらまし 認識, 思考, 情緒, 行動, および発話過程を統合的にシミュレーションする, “イソップワールド” プロジェクトを推進している. 本稿では, 1人の登場人物をシミュレートする従来の“k-システム”に加えて, 新たにもう1人の“p-システム”を構築し, 両者の間で対話を可能にすることを考える. 対話の話題としては“勧誘”を取り上げ, 勧誘する側とされる側の心理構造を詳しく調べる. 各心理構造は, 話題に応じたいくつかの要素からなり, それらがゆるく結合している. そして変化に富む対話の進行過程を柔軟に受け止めることを考慮して, 1つの例を解析する. 最後に, “マルチエージェントシステム”を目指したインプリメンテーションについても触れる.

和文キーワード イソップワールド, 統合的知能, 心理モデル, 対話構造, 言語と心の対応

AESOPWORLD - - - Toward Dialog Processing

Shoji Sugino, Hiroki Murakami, Kazuya Mera,
Masato Tokuhisa, and Naoyuki Okada

Department of Artificial Intelligence
Kyushu Institute of technology

Iizuka 820, Japan

Abstract Project AESOPWORLD is ongoing at Kyushu Institute of Technology. It attacks the integration of cognition, thought, emotions, motion, and utterance, taking an example to the character or protagonist of an Aesop fable. This report discusses its extension to two characters, who make it possible to have dialog each other. The topic is focused on “invitation,” and its structure is investigated in detail on both inviting and invited sides.

英文 key words AESOPWORLD, integrated intelligence, dialog structure, language-associating-with-mind

1 まえがき

対話理解を深いレベルで計算機処理しようとすると、人間の心理モデルやいわゆる“百科辞典的知識”が問題となる。従来の言語学における意味論は、この問題にあまり立ち入らなかった。例えばチョムスキーは、文の深層構造については議論をしたが、非言語的な知識への対応付けについては別の問題(これを“language associating with mind”と呼ぼう)と見なし、立ち入るのを避けた[1]。心理世界が森羅万象に関わり、それ故に複雑膨大である以上、止むを得ないといえよう。しかし人工知能(AI)の立場からは、この態度は窮屈である。例えばシャンクは、文の意味表現として概念依存法を提案し[5]、スクリプトやプランへと発展させた。このような取り組みにより言語と概念の間の連続性がかなり解明された。しかしAI全体の課題として、記号系のみで心的過程/データをとらえる方式は、知覚や運動などアナログデータを伴う処理との間に溝を生じた。いわゆる、symbol grounding in perception/motion と呼ばれる問題である。ニューラルネットなどの導入は、この問題への新しい取り組みとも解釈できよう。

一方、画像や音声など“マルチメディア”もしくは“マルチモーダル”な言語処理が叫ばれており、この問題について現在あちこちで活発な取り組みが行われている。また従来のAIでは知性中心の研究が大半であったが、筆者は感性、特に情緒の重要性について指摘してきた[3]。特に対話処理では、相手の気持ちや情感を考慮してシステムが発話することも求められる。

以上を考慮にいれて、我々はこの数年来“イソップワールド”プロジェクトを推進してきた[4]。本プロジェクトでは、次の4種類の“統合”を意図している。

- (1) Language associating with mind
- (2) Symbol grounding in perception or motion
- (3) マルチメディアの統合
- (4) 知性と感性の統合

従来のイソップワールドでは、寓話の世界を素材にして1人の主人公の心的過程を考察し、それに基づいて認識、思考、情緒、行動、発話などを統合的に処理するシステム(k-システム)の

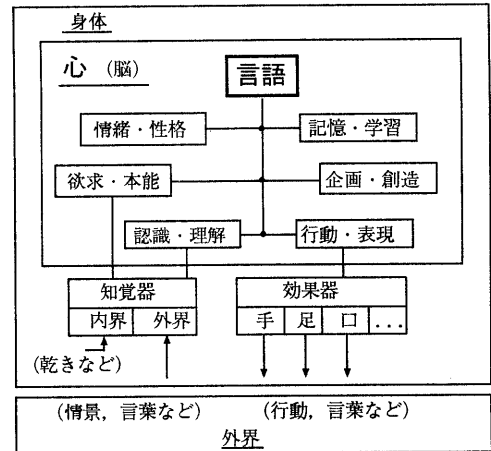


図 1: 心のモデル

構築を進めてきた。これを発展させ、現在もう1人の登場人物(p-システム)を設定し、両者の間で対話や協調的行動を行うことを目指して、k-システムの改良ならびにp-システムの構築を行っている。本稿は、これらの改良ならびに構築を上記(1)の部分に焦点を当てて報告するものである。

2 イソップワールドの概要

イソップワールドでは、寓話の主人公をシステム(エージェント)として実現しようとしている。具体的には、主人公が自然環境の中に住み、願望即ち目標をもち、それを実現すべくプランを立て、周囲を認識し、行動し、情感をもち、そして一連の心理的および物理的過程を言葉で表現するものである。以下では既に関与したk-システムの概要を述べよう。

2.1 心理モデル

まず、主人公の心理モデルを、図1に示すような9つの領域で実現しよう。幾つかの領域は、知的処理に関係している。ある外界刺激が知覚器(DS)で受理されると、その内容が認識・理解領域(DRU)で意味理解される。その結果、刺激に反応するためのプランが企画・創造領域(DPC)で練られる。続いてそのプランを実行するための行動や表現が行動・表現領域(DAE)で組み立

てられ、それが効果器 (DA) に送られて、実際の運動が行われる。一方、領域のあるものは、欲求や情緒の処理に関係している。図 1 の DS には、内界からの刺激を受理する、もう 1 つの入力チャンネルがある。ここで“乾き”などの内界刺激が知覚されると、欲求・本能領域 (DDI) で、“何か飲みたい”という欲求 (目標) が生起する。これを実現するため DPC でプランが練られ、以下 DAE, DA の順で作業が進められる。当然その過程で、前記知覚の処理過程も組み込まれるであろう。また、情緒・性格領域 (DEC) は、他の領域から情報を受け取りその価値を判断する、という重要な役割を果たす。一言でいえば、快いことは価値が高く、不快なことは低い。

記憶・学習領域は、自身をも含めたすべての領域からデータを受け取り、すべての領域にデータを供給する。言語には幾つもの側面があるが、このモデルでは、“言語の本質的役割を”心の中のすべてのデータや処理を記述する”機能にあると考えている。

次に、領域の中を少し詳しく見てみよう。心の中のデータや処理の単位を“モジュール”と呼ぶ。モジュールは、ミンスキーの提案した“エージェント”とほぼ同じである [2]。心の中には膨大な数のモジュールがある、と仮定する。各々は単純な機能しか有しないが、幾つか集まってまとまった機能を発揮する。さらにそれらが集まってより大きな機能を持ち、図 1 の各領域を構成する。

フレーム構造で表現した、モジュールの具体例を示そう。寓話の主人公の知識であることに注意して欲しい。

プランモジュール ([p 水を探して飲む,
格フレーム (立案する ([agt,\$ 自身],[obj,[水
を探して飲む])),
得る ([agt,\$ 自身],[obj,Kekka])),
目標 (程度 (Kawaki), 以下 (Sikiiti)),
具体化 ([
タイプ-1(もし周りに水があれば、それを飲む),
タイプ-2([[もし \$ 住処にいるなら,
p\$ 水瓶 1 の \$ 水 1 を飲む],
[もし \$ 野原にいるなら,
[[p\$ 住処に帰って,p\$ 水瓶 1 の
\$ 水 1 を飲む],
[p\$ 橋の下に行って,p\$ 小川の
\$ 水 5 を飲む]]]),

[もし \$ 山にいるなら,p\$ 清水に
行って、それを飲む]),
タイプ-3(i 検索 ([容器,水]),
i 検索 ([場所,水]))])

これは、p 水を探して飲むというプランモジュールである。格フレームスロットは次項で述べよう。目標は、乾きの度合いをある値以下にすることである。具体化のスロットは、3 タイプに分けて立案することを示している。タイプ 1 は、取りあえず周囲に水が見出されるならそれを飲む、という単純条件チェックの方式である。タイプ 2 は、日頃馴染んだ常套手段である。個々の手段は、それ自身 1 つのプランモジュールを別に構成しており、p 水を探して飲むと親子の関係にある。タイプ 3 は、以上がうまくいかなかったときの、深い推論による方式である。

2.2 心的過程

物語“きつねとぶどう”に沿って、主人公の心的過程とそれに伴う物理的過程を示そう。

初期状態 8 月下旬の正午過ぎ。よく晴れて、暑い。主人公は、交差点の南 300m の点にいる。とても喉が乾いているが、疲れはそれほどでもない。

欲求過程 1 欲求モジュール“乾き”が、乾きの生理状態がしきい値を越えたことを知り、DPC に喉の乾きを潤すプランニングを依頼する。条件は、“30 分以内で、危険を伴わない”、である。

企画過程 1 DPC では、p 水を探して飲むプランが起動される。p 水を探して飲むは、取りあえずタイプ 1 のプランの前提条件のチェックを DRU と DAE に依頼する。

認識行動過程 1 DRU と DAE は、それぞれ DS と DA と協力して周囲を見渡すが、水は見当たらない。

企画過程 2 そこで p 水を探して飲むは、タイプ 2 のサブプランに焦点を当てる。現在野原にいるという条件のもとで、2 つのサブプランが候補に上る。そこでこれらを評価モジュール (紙面の都合で省略) に依頼して選択しようとするが、評価モジュールからいずれも時間の不足で実行不可能の返事を得る。

次に p 水を探して飲むは、タイプ 3 の i 検索を起動し、まず水を蓄えた容器を検索する。例えば [\$ 水瓶 1,\$ 住処] が見つかり、そこへ行くまでの時間が評価されるが、条件を満たさない。検索されたすべての容器についても同様である。次に水を蓄えた場所を検索するが同じ結果である。

結局 p 水を探して飲むは、条件を満たすプランが見つからない、と DDI に回答する。

欲求過程2 “乾き”モジュールは、さらに乾きの状態が厳しい状況になっているので、条件を“多少危険であっても、急いで”に変更し、再度立案を依頼する。

DDIでは、最終的に“p\$ 館の庭へ行く、p 池を探す、p その水を飲む”という複合プラン C1 を立案する。主人公は、館に到着し中に入る。池を探すが見つかからない。館へ行って水瓶を探そうと、計画を変更する。途中ぶどう棚を発見する。何度も飛びつくが、後少しでとどかない。情緒“意地”がさらに飛びつかせるが、欲求“疲れ”が中断させてしまう。最後に“悔しい”が、負け惜しみを言わせる。

2.3 心的過程からの言語生成

非言語的な心的過程と言語の深層構造の対応付けを生成の立場から考えよう。よく知られているように、文の内容は核になる“命題情報”とそれを取り巻く“法情報”から構成されているが、ここでは前者に注目する。種々のモジュールの中でも、特に行動やプランなど動的なモジュールが“活性化”したとき、それを記述しよう。活性化の記述には幾つかの観点が有り得るが、特にモジュールの振る舞いと、それによって生じる結果に焦点を当てよう。例えば、p 水を探して飲む場合、次のような記述が考えられる。振る舞い：立案する ([agt,\$ 自身],[obj],[水を探して飲む])

結果：得る ([agt,\$ 自身],[obj],Kekka)

ここで変数 Kekka には、プランニング終了後、例えば複合プラン C1 が代入されている。

動的なモジュールには、このような“深層格フレーム”のスロットを準備しておく。そして動的モジュールの活性化に伴って幾つもの静的モジュールが活性化したとき、それらは格要素として格フレームの中に埋め込まれる。

以上の過程を言語領域が観察することにより、心的過程と言語の対応付けが可能となる。前項に沿って生成された文章の例を示そう。

とても暑い日だ。今、交差点から 300m 手前の飲み場にいる。

とても喉が乾いている。喉の乾きを潤したい。早く潤したい。危険なことをせずに潤したい。

水を探して飲もうかなあ。もし周りに水があれば、それを飲もう。水がない。今、野原にいる。住処に帰

ろうかなあ。住処は遠い。住処に帰るのはやめた。橋の下に行こうかなあ。橋の下は遠い。橋の下に行くのはやめた。なにか他の方法を考えよう。

水のある場所を探そうかなあ。池があった。池を探そうかなあ。池には B 池があった。B 池はインソップワールドにあった。B 池の近くには猟師小屋があった。多分猟師小屋には猟師がいるだろう。猟師は危険だ。B 池に行くのは止めた。

水分の多い食物を食べようかなあ。果物を探して食べよう。いちごを食べようかなあ。いちごは春に実った。今夏だ。いちごを食べる季節ではない。いちごを食べるのは止めた。

3 対話の枠組み

従来開発してきた k-システム (きつねの“コン吉”)の他に、新しく p-システム (たぬきの“ボン太”)を構築し、両者の間で対話を可能にしよう。

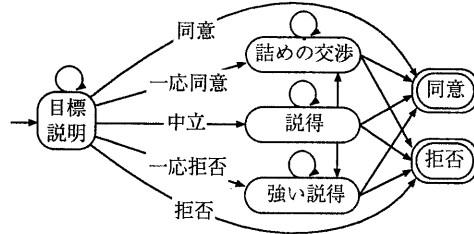
3.1 対話の環境

大きく3つのシステムから成る。

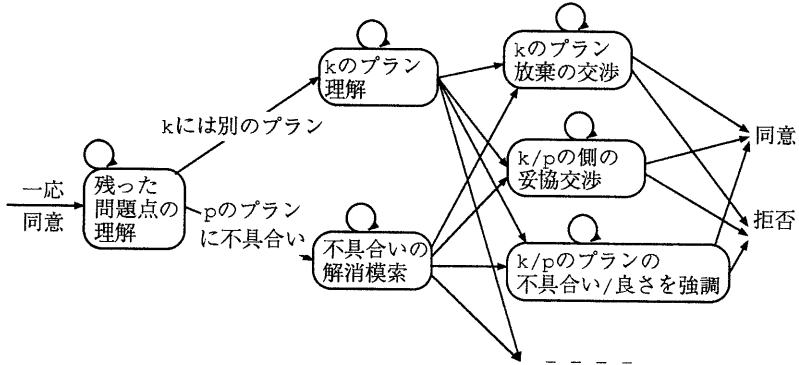
- (1) k-システム
- (2) p-システム
- (3) 外界システム

(1)は既に開発しているが、言語の解析部と次項で述べる相手モデル部が不足しており、これを補わなければならない。また(2)は新たに開発するものであるが、基本的にはk-システムの複製から出発する。そしてコン吉に関する各種のデータ、例えば記憶しているエピソードや形成している概念を新規に加えたり、情緒のパラメータを変更したりして、ボン太の個性を持たせる。最後に(3)は、感覚器や効果器を備えた自律型移動ロボットの開発を計画しているが、現時点で両システムを実際に外界で動かすことはかなわない。従ってワークステーション上でシミュレーションするには、k-およびp-システムに対し、外界として振る舞う支援システムを作成しなければならない。これについては4節でふれる。

以上の3システムが、次のように動作する状況を設定しよう：

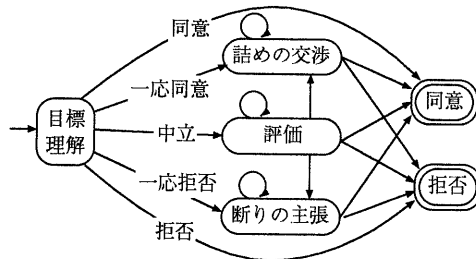


(a) 粗い推移

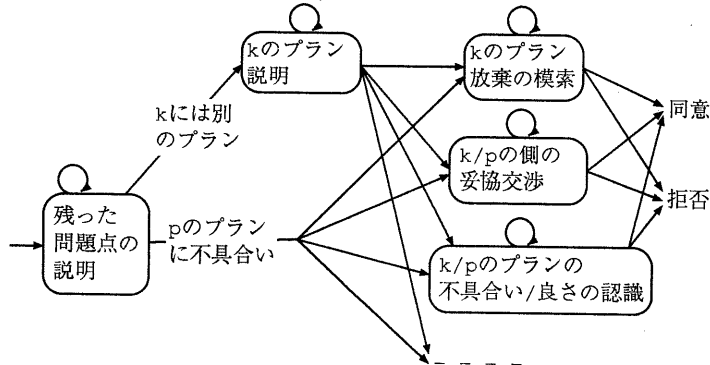


(b) 細かい推移 --- "詰めの交渉"

図 2: 勧誘する側 (p) の心理構造



(a) 粗い推移



(b) 細かい推移 --- "詰めの交渉"

図 3: 勧誘される側 (k) の心理構造

2.2 項で述べたコン吉に対し、同じ頃ボン太は館の付近を通って交差点に向かう。両者は、交差点付近で出会い、次の対話を交わす¹。

コン吉：「ボン太君、今日は。」
 ボン太：「今日は。」
 コン吉：「どこに行くの?」
 ボン太：「橋の下に魚を捕りに行くんだ。」
 コン吉：「ふうん。」
 ボン太：「コン吉君は?」
 コン吉：「喉がカラカラなので、館の池に水を飲み
 に。」
 ボン太：「館は危ないよ。」
 コン吉：「どうして?」
 ボン太：「さっきあそこを通ったとき、大勢の人の
 声が聞こえたよ。」
 コン吉：「本当? 困ったなあ。行くのをよそかな
 あ」
 ボン太：「一緒に橋の下に行かない?」
 コン吉：「えー。橋は遠いよ。」
 ボン太：「でも、あそこの水は、冷たくておいしい
 よ。」
 コン吉：「じゃあ、行こう。」
 この後、コン吉とボン太は、橋の下に向かう。

3.2 対話の構造

対話には、大きく2つのタイプがある。1つは“ゆるやか”に構造化されたもの、もう1つは比較的自由に展開するものである。構造化タイプとしては、勧誘、買い物、(医師の)診断などの話題があり、自由展開タイプとしては、(単なる)おしゃべりが典型的で、この場合は局所的な関係の連鎖で対話が進行する。以下では、構造化タイプの、特に“勧誘”に注目しよう。

勧誘する側(p)から構造を眺めると、要素としてまず、“(勧誘の)目標説明”があげられる。それに対し勧誘される側(k)の反応は、大別して(直ちに)同意、一応同意、中立、一応拒否、および(直ちに)拒否に分けられよう。その内一応同意に対しては、“詰めの交渉”、中立に対しては、“説得”、また一応拒否に対しては“強い説得”という要素がそれぞれ有り得る。以上により、図2(a)の構造が得られる。

例として詰めの交渉をもう少し詳しく眺めてみよう。100%の同意でないので、若干問題が

¹ 実対話ではなく、シナリオを意識して作成した、疑似対話。

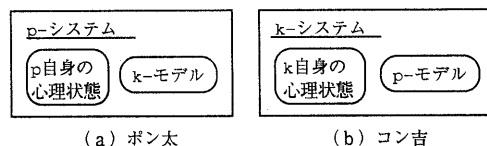


図4: ボン太およびコン吉の心理構造

残っている。そこでまず、“残った問題点の理解”という要素がある。問題点の所在を大きく、kの側に別のプランがあって不具合が生じている場合と、取り立ててプランがあるわけではないが、pのプランにkから見て不具合がある場合に分けてみよう。それぞれに対し“kのプランの理解”、“不具合の解消案模索”という要素が、交渉や説得の前段階として必要になる。そして最後は、戦略として、“kのプラン放棄”を促したり、一方が条件を変更するなど“k/pの側の妥協交渉”を行ったり、比較して“k/pのプランの短所/長所を強調”したりして、交渉をpに有利に導こうとする。以上により、図2(b)を得る。

一方勧誘される側から見た構造もある。図3(a)、(b)に示しているように、それぞれ図2(a)、(b)にはほぼ対応している。しかし、図2(a)の、p固有の要素(不具合の解消案模索)に対しては、図3(b)で対応要素がない。

図2と3により、対話者間の構造が得られた。図4にそれを示そう。p-システムには、p自身の心理状態の推移を表す構造(図2)と、相手の心理構造のモデル(k-モデル)とがある。一方k-システムには、k自身の心理状態(図3)と相手の心理構造のモデル(k-モデル)とがある。個人によって形成概念は異なるので、一般には、pの心理構造とp-モデルは等しくない。しかしここでは一致するものと仮定して議論を進める。kの心理構造とk-モデルについても同様である。

3.3 対話の進行

図2と3の矢印は、一応対話の進行を示している。しかしそれは標準的な推移、いわば推移のデフォルト値といってよい。実場面におけるほとんどの対話において、そのような推移は期待できないであろう。3.1 項の対話例に沿って

それを考察する。なお、対話構造として勧誘以外に“挨拶”も仮定しておこう。“今日は”などに始まって、天気や健康状態、さらには現在実行中の事柄などについて話して、“さよなら”を告げるものである。

p- システム

まず、ボン太(p- システム)の側から眺めてみる。ボン太は、コン吉を誘う積もりで挨拶を交わすものとする。

コン吉：「ボン太君、今日は。」

ボン太：「今日は。」

“挨拶”構造の導入部である。

コン吉：「どこに行くの?」

挨拶構造の展開部であるが、ここで勧誘構造の図2(a)目標説明部にスイッチする。

ボン太：「橋の下に魚を捕りに行くんだ。」

コン吉：「ふうん。」

交渉に入る前に相手の様子を伺う。

ボン太：「コン吉君は?」

コン吉：「喉がカラカラなので、館の池に水を飲み」

に。」

相手に別のプランがあることを知る。自分のプランに対する反応は不明であるが、一応同意が得られることを期待して詰めの交渉、図2(a)“kのプラン理解”まで進んだことにする。ただし“残った問題点の理解”は未処理なので保留である。次に、いずれの戦略で交渉するかの前処理に入る。エピソードデータを参照することにより、館の危険を想起こす。そこで“kのプランの不具合を強調”する戦略を選ぶ。

ボン太：「館は危ないよ。」

コン吉：「どうして?」

ボン太：「さっきあそこを通ったとき、大勢の人数が聞こえたよ。」

コン吉：「本当? 困ったなあ。行くのをよそうかなあ」

戦略は成功しそうなので、同意を求める。

ボン太：「一緒に橋の下に行かない?」

コン吉：「えー。橋は遠いよ。」

橋の遠さが問題であることがわかる。バックして“残った問題点の理解”処理をした上で、交渉を続行する。すなわち“pのプランの良さを強調”する戦略にスイッチする。

ボン太：「でも、あその水は、冷たくておいしいよ。」

コン吉：「じゃあ、行こう。」

“同意”が得られた。

k- システム

次に、コン吉の側から眺めてみよう。当然コン吉は勧誘を受けることは知らない。ボン太を見て、“挨拶”構造の導入部に入る。

コン吉：「ボン太君、今日は。」

ボン太：「今日は。」

コン吉：「どこに行くの?」

ボン太：「橋の下に魚を捕りに行くんだ。」

コン吉：「ふうん。」

ボン太：「コン吉君は?」

コン吉：「喉がカラカラなので、館の池に水を飲み」

に。」

ボン太：「館は危ないよ。」

ここまでは、挨拶構造が展開される。しかしC1すなわちp\$館の庭へ行く、p池を探す、pその水を飲むのプランニング過程で館の危険が検討されたいきさつがあるので、再プランニングが必要となった。

コン吉：「どうして?」

ボン太：「さっきあそこを通ったとき、大勢の人数が聞こえたよ。」

C1の実行は、ほぼ断念せざるを得ない。当然、喉の乾きを潤すための、新たなプランが必要になった。

コン吉：「本当? 困ったなあ。行くのをよそうかなあ」

ボン太：「一緒に橋の下に行かない?」

ここで勧誘構造にスイッチする。図3(a)の目標理解は終了しており、一応同意する方向なので、同図(b)の“問題点説明”に入る。k自身のプランは既に諦めており、pのプランの不具合を指摘する。

コン吉：「えー。橋は遠いよ。」

ボン太：「でも、あその水は、冷たくておいしいよ。」

“pのプランの良さを認識”したので、“同意”する。

コン吉：「じゃあ、行こう。」

次に、相手モデル、即ちk-およびp-モデルにおける推移であるが、それぞれ上記p-およびk-システムにおける推移とほぼ対応している。ただし両モデルにおける推移は相手の予測なので、自分自身の推移に比べ一般に粗いものにならない。

言うまでもないが、両モデルは相手の心を“読む”ことに相当する。例えばポン太が館の危険で動揺し、再プランニングを考えているタイミングで、「一緒に行かない?」と同意を求めるには、自身の心とk-モデルとの協調処理が不可欠といえよう。

4 インプリメンテーション

k-システム、p-システムおよび外界システムのインプリメンテーションについて概説する。

k-システムについては、2節からも理解されるように、“マルチエージェントシステム”を目指している。あるモジュールの活性化が近隣のモジュールに派生し、“活性化の連鎖”が起こる。このように自律的に動作するモジュールがグループを形成し、次第に大きな作業を行う。2.1項のp水を探して飲むの例で示したように、小さなサブプラン(タイプ1や2など)が集まって大きなプランになっている。このように幾つもの条件を示して関係モジュールを結びつけることにより、3.2項で述べたゆるやかな構造と、3.3項で述べた臨機応変の推移を実現したい。

しかし、注意して欲しい。本稿のマルチエージェントシステムでは、すべての管理が平等に分散することを意図していない。k-システムでは、“マネージャー”というグループも幾つか存在する。

k-システムの改良とp-システムの構築は、上記と同様の手法で行われる。現在、図2と3に沿って知識データを作成するとともに、信念管理のためにTMSを改良しつつある(2.3項の生成例の最後の段落に注意)。

また外界システムは、大きく、外界データベース、画像処理部、音声処理部、および行動処理部からなる。外界データベースは、木や車や道路など自然環境の他、例えばコン吉から見ると、ポン太も含む。雨や風などの自然現象、重力による落下、車の移動など、時間による変化も扱う。画像処理部は、外界データベースによりコン吉やポン太から見た2次元画像を生成する。音声処理部は、発生した音や声の伝達を行う。

以上のインプリメンテーションが現在進行中である。

5 むすび

従来から進めているイソップワールドプロジェクトをさらに推し進め、登場人物を1人から2人に増やすことにより、対話処理を可能にすることを考えた。そのため、既に関連しているk-システムを改良し、新たにk-システムを構築し、そしてそれらを支援する外界システムを作ること提案した。

特に対話の話題として勧誘を取り上げ、詳細な構造を考察するとともに、イソップワールドプロジェクトの中に自然にとけ込むインプリメンテーションを示した。対話構造については、自身の心理状態の推移だけでなく、相手モデルについても力を入れた。

複数の登場人物の出現により、イソップワールドが大きく前進し、対話における language-associating-with-mind 問題に関して1つのひな型を示すものと、期待される。

謝辞

イソップワールドプロジェクトの推進には、九州工業大学中村順一助教および乃万司助教授に、理論面でもシステム構築面でもご協力願った。記して謝意を表する。

参考文献

- [1] Chomsky,N: Aspects of the Theory of Syntax, p.251, MIT Press(1965).
- [2] Minsky,M.: The Society of Mind,p.339, Simon and Schuster(1985).
- [3] 岡田直之: 自然言語理解には情緒的機能も必要, 電子情報通信学会誌, Vol.70, No.9, pp.897-902(1987年9月).
- [4] Okada,N. and T.Endo:Story Generation Based on Dynamics of the Mind, Computational Intelligence, Vol.8, No.1, pp.123-160(1992).
- [5] R.Schank: Conceptual Dependency; A Formalism of Natural Language Understanding, Cognitive Psychology, Vol.3, No.4, p.552 (1972).