

特集「コモンセンス」

豊かな思考をもったコモンセンスシステムの構築

Constructing Resourceful Commonsense Systems

Heikki Ruuska Minsky Institute for Artificial Intelligence.
Heikki.Ruuska@sci.fi, <http://www.minskyinstitute.org>

Keywords: commonsense reasoning, cognitive systems, knowledge representation.

1. はじめに

Minsky Institute では、Marvin Minsky 著「ミンスキー博士の脳の探検」(英: *The Emotion Machine*, 以降略して「脳の探検」[Minsky 06]) に論じられているさまざまなアイデアをもとに、コモンセンスシステムの検討・開発が行われてきた。ミンスキー教授が亡き現在、Institute での活動は落ち着きを見せながらも、難関を乗り越えて研究開発を進めている。著者がミンスキー博士と長年の共同研究を行ったこともあり、本稿では「脳の探検」をシステム化した過程でもととなったアイデアを凝縮し、改善を重ねた新しい理論の一部について紹介する。とりわけ、機知をもったコモンセンスシステムの構築が目的で、そのために自己分析機能が役立つ。学習段階で、さまざまな問題に適した解決方法を模索・評価し、最適化された方法とフォールバック方法を記録する。そうすれば、新たな場面での特徴に見合った最適そうな方法が失敗しても、別の方法で試みる事ができて、これで成功したら最初の方法と異なった差分をまた記録する。モデルの基本となっているのは「脳の探検」で紹介された6重の層状構造の洗練版で、基本的な考え方は本稿でも紹介するが、より詳しくは「脳の探検」の5.1～5.5節、6.3節と7.4節を参照されたい。

1.1 コモンセンスの広義

子供は大きく分けて3種類の知をもつ。先天知、自分で周辺とやり取りをして獲得する知、そして親、メディアなどから伝わった文化にまつわる知。危険を未然に防ぐ方法、学習方法とさまざまな知識、社会的にやるべきこととそうでないことを学ぶ。日本語の〈常識〉では同一文化圏内で伝わる知の集合体を呼ぶ傾向がある。ほとんどの人間は10歳までにだいたい身に付けている。

本稿ではもっと広義として、子供が生まれながらもつ知識と、環境とのやり取りから身に付ける知識、そし

て親がもつ常識の和集合を「コモンセンス」と呼ぶ。

これほどの広義として捉えると、応用分野は多い。注意すべきところは、チェス、囲碁のスキル、運転技術そのものはこのコモンセンスの広義に入らない。しかし、それらの技術を身に付けることはほとんどの人ができて、身に付ける「方法」はコモンセンスに含まれる。

2. モデルと設計

本章では、1.1節の広義の一部に対応するコモンセンスシステムを紹介する。ただし、技術的な詳細はここで深く触れない。[Ruuska 16]を参考にされたい。その代わりに、「脳の探検」出版後の洗練・蓄積された行動モデルとそのシステム化について述べる。

探検的システムのため、仕様を決めるより、いくつかの方針と目的機能を設定する開発体系にした。シンボルと数字による方法だが、まず難点を取り上げ、次に行動モデルと知識表現を紹介し、設計の要点を押さえる。2.4節と2.5節では、本システムがどのように学習と多様性をもった思考に対応するかについて示す。

2.1 設計上の課題

ほとんどの人工知能システムは、特定の種類の問題を解くために設計されている。IBMのワトソンが「ジェパディ!」のために[Ferrucci 10]、幅広い知識を利用してはいたものの、システムの用途は限られていた。ニューラルネットワークでも、適切な目的スキーマを設定することが大前提で、そのスキーマ以外には対応できない。

機知をもったコモンセンスシステムに求められることは、多彩な種類の場面に対応できること、行き詰まったら別な方法を試みる事、失敗したら失敗から学べる事などがある。さらに、そのシステムが他のアクタとコミュニケーションが取れるために、自分の行動と目的を表現し、相手の行動と目的も表現しなければならない。コミュニケーションが必要になる場面には無数のバリ

エーションがあるため、特徴的な場面を表現するための抽象化された知識表現が必要となる。ただ、知識表現を使うシステムの難点として、情報量と人手による入力がある。そこで、ここでは学習とプロセスの学習、誤りからの学習とパターンと抽象化によって、難点を補作しようとする。ある程度の知識と複数の学習方法、そして学習方法を改良するためのプロセスが組み込まれ、あとは入出力環境とのやり取りから学ぶことが目的である。

システムを設計するうえでは、次のような課題がある。

- 〈適切な知識表現〉十分な行動要素を表現するための適切な粒度の知識表現が必要。
- 〈プロセス〉行動に至る知的活動を行うアルゴリズム。ここでは特に問題検出、分析、類似性評価、プランニング、失敗対策、学習するための評価を指す。
- 〈入出力〉直接シンボルを入力するのは簡単だが、実世界で使うときはグラウンディングの問題が残る。ビッグデータとやり取りをするときは、インタラクションのモデリングが難しくなる。
- 〈情報量〉システムが使える知識表現で記録された十分な知識量を得ることが昔からコモンセンスシステムのネックである。

2・2節ではプロセスのもとになる行動モデルを紹介し、2・3節ではそのプロセスが使える知識表現について述べる。入出力は実世界の物理的ルールと簡略英語からなる実験用データベースと簡単なシミュレーションだが、本稿では詳しく述べない。情報量の獲得に関しては、学習プロセスで補足する目的で、学習プロセスに関しては2・4節で述べる。

2・2 行動モデル

提案するモデルでは、行動がさまざまな目標（以下、ゴール）を達成するために行われる。一つ一つのゴールは、目的状況を指しているだけのものではない。活性化条件があって、達成条件があって、何かの処理を行うプログラムである（図1）。心にはたくさんの現状と目的状況との差分点があり、これらが常に〈批評家〉（クリティック）というプロセスによって検出されている。批評家はつまりゴールの活性パターンである。批評家には「寒すぎる」のような生物的機能を維持するものもあれば、「解こうとしている問題が難しすぎる」、「予想外の

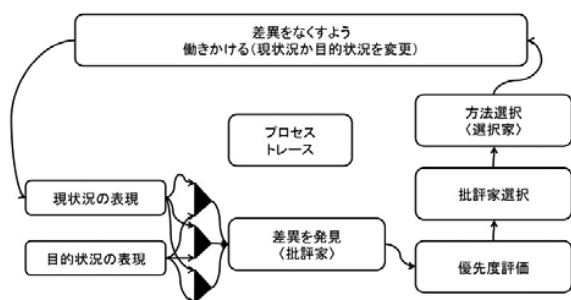


図1 差分検出による行動モデル

結果になった」といった心のプロセスの異変を発見するものもある。

批評家が起こした現状と目的状況の差分の優先度が評価され、これをもとに〈選択家〉（セレクトア）が方法を選ぶ。選択された方法の種類によって、それが直接に実行されるか、方法候補として現在状況に追加される。つまり、選択家が批評家のパターンに反応する。方法候補があると、また別の選択家がそれで活性化され、全状況に合わせて方法の結果予測、副作用分析などが行われる。

選択肢による評価の段階で、行動候補の組合せが階層構造を形成する（図2）。熱いものを触ったときのように反射的な対応は直接に行動をつくる。車道を渡ろうとするときは、行動結果を予測するために〈予測行動候補〉をつくる。こんな行動が選択家によって選ばれると計画的な熟考行為が始まる。予測ができれば、それも予測として現状の表現に入る。無害な予測と行動候補に反応する批評家がまた選択家を起こし、行動がようやく実行される。そして行動を取って、予想外の事態が起こったときに、予想プロセスを分析し、間違った考えを探そうとし、今後のためにどうすればよいかを検討する。これは内省的階層の行為である。こうやって行動に至るゴール活動の性質から階層構造が成り立つ。ほかにも、例えば同じことを繰り返しても、目的状況に近づかない、結果を予測しようとしても、熟考するのに時間がかかりすぎるなど、プロセスの性質に反応する批評家もある。それぞれの階層の機能については「脳の探検」の5章で詳しく述べられている。心の批評家-選択家モデルについては同書7.2節で説明されている。モデルに追加した学習ゴールについては、本稿の2・4節で述べる。

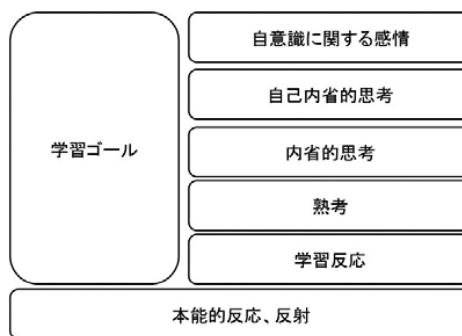


図2 心の階層モデル

2・3 データモデル

データ構造は類似問題検出性とプロセスの柔軟性を重視している、シンボルと数値からなるダイナミックウェイトグラフ（Dynamic Weight Graph）を使っている。これは問題解決に有効な類似性検出、問題分解、観点変更を重視して選んだ。

人間は多くの問題解決に思考がいらぬ。これは、すでに解決方法を知っているからである。解決方法がわか

らないとき、似た問題の解決方法を知っていれば、それに適切な置換をして応用できるときがある。例えば、目覚まし機能の付いた携帯電話を忘れていても、ホテルの目覚まし時計が同じ目的に使える。

また、問題分解を可能にするため、問題をアクティブにした要素を別々に表現できる必要がある。例えば、寝坊して会議に遅れそうなとき、服装、持ち物などの準備と移動手段の双方を考慮しなければならない。「会議に参加する」ことの前提要素と現実の差分として現在位置といくつかの準備不足があるとすると、それぞれに対して別々にプランニングできる。

さらに、観点変換について述べる。前例の寝坊者が会議場へ走って急ぐ道に、長い下り坂がある。そこで坂上のスーパーマーケットからショッピングカート借りてこれで坂を下る。この場合は、カートは、物を運ぶ道具から、車輪の付いた移動手段に再表現された。具体的には、構造的関連上で類似性を発見し、新しい機能的なエッジが挿入される。もしその者が後で転んだり怖い思いでもしたら、このエッジを批判する批評家ができ、その批判理由は操作性のなさとなるだろう。

このような場面に対応できるのは、自然言語に準じたシンボルでつくるグラフだが、エキスパートシステムのオントロジーと違って、論理的一貫性が求められていない。データでの異変と矛盾は批評家によって処理される。例えば、「すべての獅子は危ない」、「ぬいぐるみのライオンちゃんは獅子である」の両方があっても、役立つ知識として同時に存在し得る (図3)。

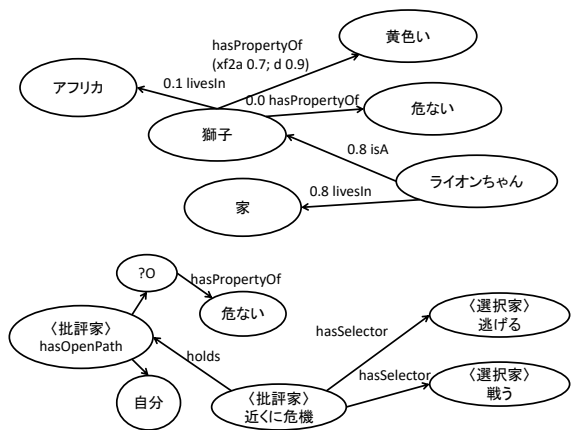


図3 データモデルのインスタンス

ライオンちゃんを扱うゴールは「危ない」へのエッジをゼロに近いウェイトにする。逆に、獅子から逃げる反射的選択肢を活性化するのに、獅子を近くから見る以外のクリティックが必要であり、そうでないと動物園で獅子を見たら逃げてしまうだろう。

つまり、古典的なオントロジーなら、獅子に危ない動物とそうでないぬいぐるみに別々なシンボルを付けて、それぞれの親存在として何かの抽象的な獅子を置くが、

本システムでは、ゴールが行うエッジ操作によってその必要性がなく、曖昧さが許容される。本システムの場合ごとのエッジのウェイトはアクセスするゴールとシステム全体の部分ハッシュで識別され、数値は2・5節で紹介する学習によって変更する。

ここまでは、2・1節で述べた課題のうち、プロセスと知識表現に該当する行動モデルとデータ構造のシステム化について説明した。入出力に関しては、現在はルール、ファクト、簡単な物理シミュレーションと簡略化された英語ベースの表現を使う。すべてがシンボル表現で、実世界とやり取りをするにはグラウンディングが必要であるが、ラベリングの研究が盛んに行われているため [Asai 17, Sharma 14], 現段階では対象としていない。システムのテクニカルな詳細は本稿で省くが、基本設計は [Singh 05], 再構築・改良版は [Ruuska 16] を参考にされたい。次に、システム化が進んで可能となった行動分析、モチベーションと学習の相関と機知をもった問題解決について述べる。

2.4 行動：学習

問題を2・2節に述べた方法で解いてから、学習が行われる。ただし解き方を知っていて、予想外の結果にならなかった場合には学習の必要がない。学習は使った解き方によって異なる。

新しいファクトを学ぶことは、「未知のファクト」と「学ぶべき」の批評家が活性化され、選択家の操作はグラフに新しいノードとエッジを入れることである。その直後の学習は、単にエッジの標準値を設定するだけである。

次に簡単なのは、矛盾の少ないセマンティック拡張、例えば「Fish breathe water」, 「Animals breathe air」に例外を入れて、「Animals except for fish breathe air」と表現を直せばいい。グラフでは魚のインスタンスが生き物枠の変数に入っているとき、部分ハッシュによって「breathe」のターゲットで「air」が0, 「water」が1となる。

ファクトを学ぶ以外の問題解決で成功した場合は、成功に至った手順が記録されるが、もっと重要なのは有効だったプロセスを正しく評価することである。そして試みたが失敗したプロセスも評価する。失敗したプロセスに対して、失敗を判断した批評家も登録する。

具体的な学習プロセスは学習ゴールによって処理される。以下の積木の例では、できるだけ高い塔を建てると

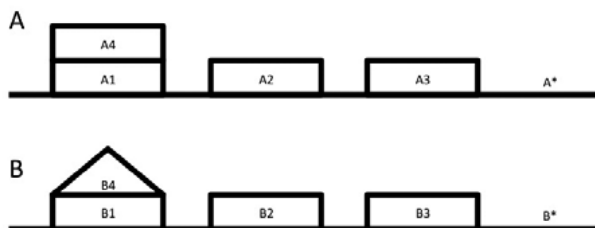


図4 積木で塔を建てる問題

いう目的状況をもっている場合を考えよう。

A の場合、積木を横 (A*) に置くことには逆効果がある。もし A4 → A*, A1 → A2, A3 → A4, A3 → A2, A4 → A2 の操作を行ったら、A4 → A* は評価すべきではない。目的状況から遠ざかったからである。一番高い塔の上に積木を置く以外の行為は目的に対してむだである。しかし、このルールを発見するのはハイレベルな学習ゴールであり、そのゴールの批評家は、ステップごとに目的状態に接近すれば、目的状態に達することを検出する。つまり、差分を少なくしさえすれば、むだなステップはなくてすむ。

しかし、A で学んだ方法を B の場面で使うと、最初のステップで失敗する。サポートが積木を置くこと的前提である。これを選択家の前提に入れると、B2 → B3, B4 → B3 ができるが、ここで終わってしまう。ここまで来たら作戦がいくつかある。例えば

1. 最も高い塔をつくったと暫定して諦める。
2. 最初の状態に戻して網羅的な検索を行う。
3. B1 が B2 と B3 に似ていると気づき、行き詰まった状態で B4 と B1 が交換できないかといういろいろ試す。

実際の子供はどの方法もやっているが、やがてさらに抽象化されたルールに気付く。すなわち、上をサポートできない積木が塔の一番上になる、というものである。モデルでは、どの行き詰まったまたは完成した状況においても、この性質の積木が塔にあるならそれが一番上にある、というパターンがあることをハイレベルなゴールが検出し、グラフに塔の属性としてそれを追加する。

以上では、ヒルクライムといった一般的な方法とその欠点を補足する一例を紹介した。他のヒルクライムによる行詰りを避けて、学習が続けられる戦略例として「刺激」をあげる。刺激を求める行為は、結果が予測しにくい問題をわざと解く。多くの批評家が警戒信号を鳴らす。ハイレベルな批評家が学習効果を評価して警戒批評家を黙らせる。リスクもあり、解いている最中は不快を訴える批評家が多いが、成功したらリワードが大きい。これを充実感と呼び、ハイレベル批評家の評価方法である。

ここで紹介したように、階層構造モデルはさまざまな学習方法に対応できる。最後に学習のモチベーションについて述べる。行為を、リワードを求めての行為と、系統的に起こされる行為 (Architectural motivation [Sloman 16]) に分けると、子供がコモンセンスを得るための行動のほとんどが後者に分類される。子供が特に褒められなくても必死に遊び、試行錯誤をし、親その他の大人にたくさんの質問をする。リワードを求めての行為とそうでない行為の学習評価に関して相違点があるかどうかは今度の課題である。

2.5 思考の豊かな行動

スキーリフトで閉店時刻を過ぎて、ゴンドラに乗った二人。スタッフは人が乗っていることに気付かずにゴンドラを止めた。今度動くのは、翌日の朝。夜はマイナス十数度と寒い。飛び降りると斜面まで十数メートル、怪我をする。二人が助かったのは、暗くなってからもっていた紙幣を燃やし、誰かがその明かりを見たからだ。

「思考の豊かさ」には複数の性質があり、ここではものを多様な観点から見ること、複数の重複した方法でゴールを達成する方法について述べる。上の実例では複雑な思考が使われた。その中でも紙幣を通貨ではなく、燃やせるものとして扱ったことが特に注意すべき点である。本稿の行動モデルでいえば、知識表現を変換したのである。もっと簡単に分析できる事例でいうと、図4の積木に戻る。2.4節の例で、積木は塔の建築材だったが、他の用途もある。もしこちらを見ていない人の注意を引きたければ、積木を投げつける。兄弟がもつぬいぐるみで遊びたかったら、積木セットと交換できるかもしれない。

これらのロール変換も本モデルで表現できる。投げる物体を探すときに、パターンマッチングで適した物理オブジェクトとして表現される。交換物を探すときに、価値あるおもちゃとして表現される。彫刻の対象としては、ほとんどの物理オブジェクトで扱えるが、金属製のものを検討する際には批評家が活性化し、エッジが低い値になる (図5)。

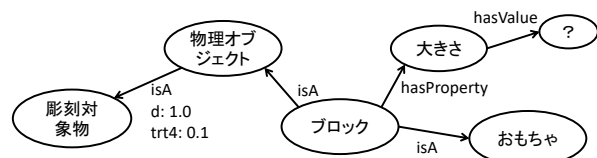


図5 データをセレクトティブにアクセスする

続いて重複した方法で同一ゴールを達成することについて述べる。「脳の探検」では、距離を計るさまざまな方法について述べている。ここでは、コンピュータにとっては簡単だが、コモンセンス的にそうシンプルではない「足し算」について紹介する。図6 (A) ~ (D) の点を数えるのにいくつかの方法がある。子供が数えるとき、同一点を2回数える誤りがよくあって、これを防ぐ方法と、タスクを早くする方法がいくつかある。

どの場合でも、数えた点に印を付ければ、2回数えることが防げる。

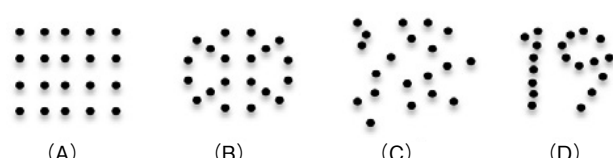


図6 点を数える問題

図6 (A) では、掛け算がわかれば、行と列を数えて掛け算する。これは最初に述べた別の観点から見る方法。これは知識表現変換を利用した。

図6 (B) では、(A) の方法は使えないが、図を上下左右四つに分けて、対称を使ってどれも同じ個数をもって、4分の1の点々を数えてまた掛け算を使う。これは問題分解と類似性検出を利用した。

図6 (C) では、掛け算は簡単には使えない。線を引いてグループ分けして、それぞれを数えて足し算か、一つ一つマークするかの方法がある。パターン検出による問題分解、あるいはブルートフォースを利用した。

図6 (D) は罫である。ここまで (A) ~ (C) を数えたなら、ハイレベルなパターンから数は同じ、数えることがむだという批評家が活性化する。だが、図を見たパターン認識が点からなる数字を検出し、別の批評家が活性化する。さらに、本稿の読者は心理学の実験を知っていて、点の数をパターンの数字でだまされていることがあるので、ハイレベルな批評家が最初のパターンをサポートするかもしれない。結局これも数えるしかない結論に至る。しかしその場合も、任意な点から数えるのを始めるのではなく、左右にグループ分けして、それぞれを線として表現し、2回数えることを防ぐ。

図6 (A) ~ (D) のすべてにはブルートフォースが使えるが、複数の思考方法を使った問題解決によって人間はもっと早くゴールに達成することができる。本モデルでの処理としては、見つかったパターンに応じて批評家が活性化し、複数の思考方法の組合せによって最適な方法を選ぶ。(A) のパターンには、(B) ~ (D) で使えるすべての方法が使えるが、初めてこのパターンと遭遇したときは知っているすべての方法で試し、特に (A) より数倍大きいグリッドで掛け算が有利なのでその選択肢を強化する。だが今度、点を数えるゴールで (C) のパターンと遭遇したら、標準の方法が使えず、(A) のときの他の方法にフォールバックして、最適なのが見つければ新しい選択肢をつくって、グリッドの批評家が活性化していないときは掛け算が無効化され、(C) の選択肢が強化される。具体的な働きは図5で説明したのと同様である。

3. 関連研究と今後の課題

本稿では、「脳の探検」のアイデアの一部を加工し、システム化するプロセスを述べ、柔軟な対応の行動例を紹介した。本章では関連研究を紹介するほか、課題がまだ尽きないほどある中、今後人工知能エージェントが自主的に発達していくのに役立つと考えられる2点を紹介する。すなわち、知識表現の獲得と普遍性のある知識の発見である。

3.1 関連研究

現在、ニューラルネットワークの研究で多くの成果

が得られた。その代表例の一つは、囲碁をマスターしたAlphaGo [Silver 16]。ただ、知識表現をもっていないAlphaGoは、聞いてもどんな作戦をもとに石を置いているか表現できない。AlphaGoの改善版は強化学習を使うが、むしろこのアプローチのほうが本稿のアプローチに近い。ルールをもとにゲームプレイをジェネレートして、シミュレーションによって囲碁の戦略を洗練させる考え方が似ている。しかし、改善版でも知識表現がなく戦略を表現できず、相手の石を囲もうとするとか、石を変なところに置いて相手をだまそうとするなどの表現ができない。ディープラーニングの表現性を向上させるためのプロジェクトがDARPAによって発動された[Gunning 16]、この研究の成果が近い将来期待できる。

古典的なシンボリックシステムの継承といえるCyC/Lucid.ai [Guha 90] が多くのコモンセンス問題を解け、学習によって知識を増やせるようになった。ただ、論理的一貫性を求めるので、本モデルのようなコンテキストに応じた解釈がない。

現在開発されている認知アーキテクチャの例として、MicroPsi [Bach 09] とMECA [Gudwin 17] がある。前者では一報から複数のゴールに応じて行動をする、他方でシンボルに反応し、シンボルのマップがつくれる点では本モデルに似ているが、トップレベルゴールが固定されているため、文化で伝わる知のモデリングはない。行き詰まったところに別の表現で対応できないところが違う。MECAはSOAR [Laird 96] に似ているが、シンボルグラウンディングができる。しかし、基本的な動きはルールベースシステムと同じく、リソースをオンオフにするなど、本モデルのコンテキストに応じた対応が実現されていない。

モデルとして最も似ているのはSlomanのCogAffとMeta-Morphogenesisプロジェクト[Sloman 16, Sloman 18]である。CogAffのアイデアの一部は2004年までSimAgentとして開発されたが、プロセスの修正・生成までは現在システム化されていない。

3.2 今後の課題

§1 知識表現の獲得

現システムの欠点の一つは、新しい知識表現を発見できないことである。基本的なプランニングと失敗から学習ができて、分類できない問題には対応できない。人間は、規則性を抽出し、適切な表現で整理しようとする。その中には、以下のケースがある。

- 既知のデータをもっている表現で扱う。これは一般的な問題解決。
- 新しいタイプのローデータを既存の表現に準じたもので扱う。これもよく使われているアナロジーである。
- 既知のデータを、新しい方法で表現する。これが起きるのはまれなようで、発明や科学的発見に至るこ

とがある。

- 新しいデータ、既存の表現が使えない。人間にとってはノイズに見えるか、特徴に気付かないので表現しないが、機械学習では特徴抽出できることがある。例えば [Wang 17] のネットワークで恋愛サイトプロフィール写真から同性愛を高確率で検出できた。

他者のやり取りを見て、ものを交換するパターンを観察した学習ゴールが、パターンを保存し、仮ラベルを付けると、図7ようになる。

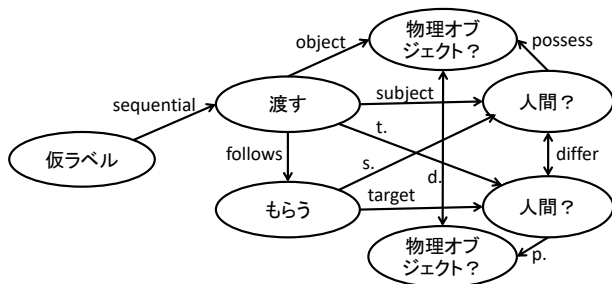


図7 連続した行動をパターン化した例

しかし、図7中の「物理オブジェクト?」と「人間?」の項目は、観察したインスタンスが共有する最大の共通分母で、さらに抽象化して、情報の交換を表現するなどには至らない。このようなコンセプトの拡張、そしてコンセプトにラベルを与えることは手動で入力しなければならない。さらに、学習したパターンでは、メタファーを使うことができない。例えば子供に、肺は血液に酸素を渡す、代わりに二酸化炭素をもらう、脳と筋肉は酸素を食べて、二酸化炭素を体から出すと説明したほうが、肺膜と溶解度平衡式と好気呼吸を説明するより理解が早い。

この知識表現の獲得を、例えば [Winston 70] のように実装したら、プログラムにもものを見せたり、またはプログラムが自発的に「人にもものをあげて、その次にその人から別のものをもらうことをたくさん見たけど、それなんていうの」と聞くようになるか、興味深い今後の課題である。

§2 普遍性のある知識の発見

一般的な事実を発見することが抽象化につながる。一般的な事実の例として、足し算の順番はどちらでもいい、閉じた箱から物を取り出すことができない、人は自分の親であり得ない、物はひもで引っ張れるが同じように押すことはできないなどがある。

図5、図6の積木に対して、塔の積木の数にかかわらず上をサポートできない積木が0か1個あり、もし1個あるなら塔の一番上にあるといった普遍的事実がある。抽象化すると、これは回帰関数とトポロジー上のターミネータの理解につながる。さらに図6の(A)の点は実世界の駒だとすれば、(B)か(C)の形になるように動かしたら、後者の数は(A)と同じと数えなくてもいい。

この1:1の関係を表現できることも人間には先天的ではなく、[Piaget 23]の実験である程度の発達過程を経てしか表現できないことがわかっている。よって、この抽象化も、学習ゴールかそれに準じたプロセスで発見・表現できるはずである。

4. 謝 辞

本稿で紹介したシステムの元になったアイデアは Marvin Minsky により、システム構築は彼と Aaron Sloman との議論によって洗練された。本稿の執筆にあたっては小侯敦士氏と桐山伸也氏に支援を受けた。

◇ 参 考 文 献 ◇

[Asai 17] Asai, M. and Fukunaga, A.: Classical planning in latent space: From unlabeled images to PDDL (and back), *Cognitum' 17 at IJCAI' 17*, Melbourne (2017)

[Bach 09] Bach J.: *Principles of Synthetic Intelligence*, Oxford University Press (2009)

[Ferrucci 10] Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., Lally, A., Murdock, J. W., Nyberg, E., Prager, J., Schlaefler, N. and Welty, C.: Building Watson: An overview of the DeepQA Project, *AI Magazine* (Fall 2010)

[Gudwin 17] Gudwin, R., Paraense, A., de Paula, S. M., Fróes, E., Gibaut, W., Castro, E., Figueiredo, V. and Raizer, K.: The multipurpose enhanced cognitive architecture (MECA), *Biologically Inspired Cognitive Architectures*, Vol. 22, pp. 20-34 (2017)

[Guha 90] Guha, R., Lenat, D., Pittman, K., Pratt, D. and Shepher, M.: CYC: A Midterm Report, *Commun. ACM*, Vol. 33, No. 8 (1990)

[Gunning 16] Gunning, D.: Explainable artificial intelligence (XAI), *DARPA Project Proposal*, DARPA-BAA-16-53 (2016)

[Laird 96] Laird, J. and Rosenbloom, P. S.: The evolution of the Soar cognitive architecture: D. M. Steier and T. M. Mitchell, (Eds.), *Mind Matters: A Tribute to Allen Newell*, pp. 1-50, Mahwah, NJ: Erlbaum (1996)

[Minsky 06] Minsky, M. 著, 竹林洋一 訳: ミンスキー博士の脳の探検—常識・感情・自己とは, 共立出版 (2009)

[Morgan 13] Morgan, B.: A substrate for accountable layered systems, PhD thesis, Massachusetts Institute of Technology (2013)

[Piaget 23] Piaget, J.: *Le Langage et la pensée chez l'enfant*, Delachaux et Niestlé, Paris (1923)

[Ruuska 16] Ruuska, H.: Critic networks for commonsense problem solving, *Cognitum' 16 at IJCAI' 16*, New York (2016)

[Sharma 14] Sharma, A., Tuzel, O. and Liu, M.-Y.: Recursive context propagation network for semantic scene labeling, *Advances in Neural Information Processing Systems*, Vol. 27 (2014)

[Silver 16] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., George van den Driessche, Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D.: Mastering the game of go with deep neural networks and tree search, *Nature*, Vol. 529 pp. 484-489 (2016)

[Singh 05] Singh, P.: EM-ONE: An architecture for reflective commonsense thinking, PhD thesis. Massachusetts Institute of Technology (2005)

[Sloman 16] Sloman, A.: The Cognition and Affect project, Website, <http://www.cs.bham.ac.uk/research/projects/cogaff/> (2016)

- [Sloman 18] Sloman, A.: The Meta-Morphogenesis project, Website, <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html> (2018)
- [Wang 17] Wang, Y. and Kosinski, M.: Deep neural networks are more accurate than humans at detecting sexual orientation from facial images, *Journal of Personality and Social Psychology (preprint)* (2017)
- [Winston 70] Winston, P. H.: Learning structural descriptions from examples, PhD thesis, Massachusetts Institute of Technology (1970)

2018年2月22日 受理

著者紹介



Heikki Ruuska

フィンランド出身。静岡大学情報学部・同大学院を修了後、ハーバード大学所属、MITのマービン・ミンスキー教授と共同研究。現在 Minsky Institute for Artificial Intelligence においてミンスキー教授のアイデアに基づいた認知アーキテクチャを設計担当。