

1. はじめに

この世界を生きる汎用人工知能の構成を考えて行って見ます。この世界はフィジカル空間です。それはイメージ（マップデータ）の世界です。従って、汎用人工知能は情報処理を統一的且つスムーズに行うために、フィジカル空間のキャッシュとして、内面にイメージの集まりである、時間と空間で記述されるイメージ（マップデータ）を持たねばなりません。それがサイバー空間です。

汎用人工知能の内面がフィジカル空間と異なるのは、記号が支配している知識空間であることです。記号によって、柔軟で、飛躍するような発想を生む素地が生まれるのです。記号空間である知識空間は知識ベースとして、構造を持ったコマンド群で記述されます。そして、記号はサイバー空間の各要素データにより意味を持つ・・・フィジカル空間に接地されるのです。

汎用人工知能は、センサーによって外界の状況を時系列的に知っていくことになります。また、アクチュエータにより外界に働きかけます。その都度、共起する記号とイメージが意味を持って、内面に取り込まれていきます。その共起する時系列データを記録しておくエピソード記憶が学習によって、世界のことがらを知るために重要なデータとなります。

結局、汎用人工知能は次の3種の空間を持つと言えました。

- (1) サイバー空間
- (2) 知識ベース
- (3) エピソード記憶空間

知識ベースはサイバー空間によってフィジカル空間に接地されます。

2. サイバー空間と記号空間（知識ベース）

コマンドは記号の内、操作・行動を表現する物で、本質的に記号の生成システムから生成されるものです。記号の生成は、センサーとアクチュエータの共起発生する（key,value）として汎用人工知能の自動的な動きとして行われます。それは「意志」を生成するシステムでもあります。共起した事象が同じ物だと、同じ記号のもとにそれら共起データは集約されなければ成りません。時系列のより前に発生した共起データの記号をもって集約のための記号とします。このようにして記号は接地されます。

特に注意しなくてはならないものにコマンドがあります。コマンドは、（動詞記号[、格記号、名詞・属性詞記号]・・・）として表現されますが、格はオントロジーとして、サイバー空間のあらかじめ認識できるものとして設定する必要があります。ですが、真っ新たな人工知能を考えて行くと、イメージの重ね合わせをデータとした、クラスタリングで得ていける物であるということです。これは十分注意して汎用人工知能を設計していくとき考慮すべきことです。

コマンドは、サイバー空間内のオブジェクトやシーンを解析して得た情報になります。それで、サイバー空間のマップのノードにコマンドは添付されて利用されます。これは、必要に応じて、汎用人工知能が生活しながら充実させていくという知識ベースに成っていきます。

コマンドは、対応する動画が基本的にサイバー空間のイメージとして存在します。コマンドが発行されると、イメージが連想され、そのイメージに添付している情報（属性などの記号や他のイメージ）を発火し、作業メモリ（空間）に展開します。その空間はいくつかのコマンドのイメージと連合され、文脈を成し、解析の対象となります。

サイバー空間は世界を一意に管理する空間が1つ在って、目標別のローカル空間があって、視界があります。それぞれ SLAM によって統一的に管理されます。

プロダクションシステムとか行動・思考は知識ベースを中心に行われる機能です。だから、サイバー空間から知識ベースへとデータが加工されていく必要があります。

3. イメージを中心にした情報処理

イメージは、時間と空間（配置）とオブジェクトとその属性によって構成されます。基本的に画像とか動画像を基本にすえています。そのなかで、オブジェクトは、画像という一つの枠組から、その構成要素を切り出して得られるものです。そうして、サイバー空間は、画像とか動画像に加えて、その解析結果をオブジェクトとして時間と空間に配置した物として構成されます。まずは、オブジェクトが配置され、そこにオブジェクトが写っている画像が配置されます。

画像からオブジェクトを切り出すときには、枠と部品群という基本的な構成が手がかかりになります。部品もその内部から見れば枠であり、枠と部品群というのは入れ子になっています。枠の切り出しの手がかりは、閉じ線の濃さです。もっとも濃い閉じ線が輪郭線に成る可能性が高いのです。それは、画素の数値の微分係数が最もたかいものを輪郭線と取るということで良いはずです。テクスチャも同様にテクスチャの輪郭線はもっとも微分係数が高い線分の分布をなぞることにより検出されるものです。

部品は、枠の中の部分的な線分の集まりです。このとき、問題になるのは、枠と部品群が1つの画像から取り出されたとして、もう一つの画像から取り出される枠と部品群のあいだに対応関係があるかを判定することです。どうやって同一人物であるか判定するとか、同じ人物が異なる服を着ていて、その服の間に対応関係は何かということですね。帽子の対応、上着の対応、ズボンの対応。靴の対応ですね。どうやって対応していき、靴なら靴にいろいろな形態があるということを経ラスタリングして認知していけるようになるのでしょうか。そもそも、枠である人物がどうして一致していると判断できるのでしょうか。

基本的に最初に得た画像（イメージ）を今回得たイメージのパターンが一致（重要事象が部分的に）していると相同と判断することになります。見た目が異なっても声が同じパターンであるとか、振る舞いが同じパターンであるとかの場合には相同と判断します。

2つのイメージの中の部品としての対応関係ですが、一つのイメージ要素の枠の部品が他のイメージ要素の枠のどの部品に対応するかという問題ですが、それは基本的に曖昧です。相対的な順序として近くにあるもの同士が対応しているという判断と、パターンが一致するとか、重要度（大きさ）が一致するとかを手がかりにした対応です。だから、これもパターンの一致が重要な手がかりに成り、対応関係が推論されていくと言えそうです。

パターンマッチングの方法として次の技術が考えられます。

- (1) 記号列マッチング
- (2) ディープラーニング

記号列マッチングは、完全一致するのはまれで、欠落後記号とか余計な記号が入り込んできたりします。

長さが異なる記号列のマッチングとして、

- (1) 左端整列
- (2) 右端整列
- (3) 中央記号整列

がありえます。いずれも、同じ記号間の距離（間に入る記号数）のガウス分布のポテンシャルをもって、重み付け一致度評価をすると良いでしょう。

ディープラーニングは、マッチさせようとするデータ群とその標準イメージを学習させ、その学習の後、出力されるイメージを対応付けたいイメージとの一致度をみることになります。

このとき、ディープラーニングの出力であるイメージとしての標準イメージを三角とか円とか四角とかの基本図形にとって学習させ、実際のイメージはこの基本イメージのセットを入力としてディープラーニングすると良いと思われま

4. エピソード空間

エピソード空間は、生活の中で得るセンサーやアクチュエータ値を土台として、関連発火しているコンセプト（オントロジーを含む）を時系列で記録している空間です。それは、主にポイントセットであり、実データはそれぞれの部分空間（センサー値空間、アクチュエータ値空間、サイバー空間、知識ベース）に存在しています。実データの削除はそれぞれの部分空間が行います。

エピソード空間は、最もサイバー空間と知識ベースが密接に連携している場です。時間マップ、空間マップを構成し、そこに今発火している事象を写します。事象を構成するオブジェクトを配置していきます。エピソード空間は、時間マップについてトコロテン式のデータ保存を行います。即ち、海馬システムになります。そこから、SLAMにより、サイバー空間と知識ベースにデータを反映させていきます。

データの反映は、記号名とかコマンドのパターンでサイバー空間と知識ベースを検索して、もっともマッチするところに組み込むことで実現します。こうしたことを考えて行きますと、汎用人工知能への情報の入口と出口はエピソード空間であると言えることが分かります。認知や行為のキャッシュであるエピソード空間が在ると同時に、エピソード空間にある事象データを整理統合してサイバー空間と知識ベースに反映させるあいだ保存しておくエピソード空間があるということです。正に海馬システムの働きそのものです。

5. おわりに

養殖マグロの個体数を数えるということを考えて見たいと思います。サイバー空間にマグロの個体を同定したオブジェクトを全部配置し、運動を予測していけば、現実のマグロの個体とのマッチングが取れて、数は正確に捉えることができるというものです。そのためには、マグロの個体を動画像より得て、サイバー空間に配置していく必要があります。

個体の画像からの切り出しは、輪郭線を発見するところから始まります。線分のもつ境界の一部らしさを、微分を行うことで得ていきます。閉じ曲線になれば、輪郭線が発見されたと判断して良いでしょう。動きは、画像のパーティクルフローを解析していけば掴めます。あとは、サイバー空間に重なるの無いように配置し、行動を予測していくことです。行動は、部分空間を作ります。全ての個体を一度に画像に納めることはできないからです。生け簀を複数設けるなどしてこれに対応していきます。

このように、サイバー空間を設けることで、情報処理が簡単になることが分かります。

RPA(Robot Process Automation)へのサイバー空間の応用を考えて見ます。

文書名とかオブジェクト名で無く、入力データそっくりそのままエピソードとして保存しサイバー空間のオブジェクトとします。そして、その上の人間による操作をコマンドとして記述するのです。オブジェクトもどういう環境のどういう記号列の流れの中のものかということ記録するのです。

ページをめくる操作とか、ページを閉じる操作とか、表中のデータを検索してオブジェクトを得ていくというような一連の操作は、イメージとコマンドで記録するのです。コマンドは、必要な情報をオペレータから得ることもあるでしょうが、殆どは自律的に記号化として獲得されていく物にします。名寄せなんかは、オペレータに確認を取った方がいいでしょうし、「リンゴの収穫」とかのオブジェクトを得るには、農産物の生産に関わる文書を検索した方がよいというような判断ですね、これは共起の解析で自律的にコマンドを設定するのが良いと思われます。

このようにシステムを設計すると、似た入力データであれば、オブジェクトを簡単に取り出せることができます。

おわり