

Artificial Intelligence in Search of the Origins of Human Intelligence (Analysis of the European Palaeolithic Cave Art)

人の知能の起源を求める人工知能

(旧石器時代ヨーロッパ洞窟壁画の解析)

André WLODARCZYK
Université Charles de Gaulle – LILLE 3
andre.wlodarczyk@univ-lille3.fr

Georges SAUVET
Université PARIS 13
sauvet@univ-paris13.fr

CNRS Action: "Origins of Man, Language and Languages"

Project : "Emergence and Function of Semiotic Systems
in Human Groups of the Upper Palaeolithic Period"

CNRS Joint Research Unit 5608

Maison de la Recherche

5, allées A. Machado, 31058 TOULOUSE – France

アブストラクト:

未知の記号システムに関する発見手順というものはまだ見つかっていませんが、フランスとスペインの旧石器時代洞窟壁画・線刻画に共通に見られる、構造上の特徴が存在するというを論証する一般的な手法を提案したいと思います。この手法は決してヨーロッパの旧石器時代洞窟芸術に特有なものではなく、他の地域の岩面芸術にもあてはまるのではないかと私達は考えています。ここでは、以下の科学分野と手法を論じます。1. 構造的記号論(意味的対立、コミュニケーションシステム)、2. 統計的データ解析(要因解析とクラスティングアルゴリズム)、3. 自然言語処理(構文解析と文法推論)、4. データベースからの知識発見(ラフ集合近似)、5. 統計的比較のアプローチ(旧石器時代岩面芸術と、アフリカおよびオーストラリアの民族的データの比較)。

1. はじめに

Leroi-Gourhan (1965) が、非常に単純な統計論を用いて論証しようとしたのが、旧石器時代(1万年から3万2千年前)の動物や非表象的壁画の(個々の)絵柄が、洞窟内(全体)での相互相関と位置を見ると、決して無原則に分布していたのではない、ということでした。ここで(私達)は、岩面芸術が集想的思考を担っているとすれば、何か別のコミュニケーションシステムとして構造化されているはずである、と主張します。

2. 構造的記号論—非表象絵柄のタイポロジと相関規則

タスク: 非表象的芸術が構成的であるかどうかを決定し、非表象的パターンすべてのタイポロジを確立する

データ: 非表象的絵画(の絵柄)の手製データベース

非表象的(あるいは「抽象的」)絵柄を、日常的に用いる物体を個別に図式表現したものとは見なさず、各絵柄の間に存在する、ある種の規則的関係を強調することにより、これらが記号システム

の一部である、と提案します。

非表象的絵柄の意味内容はすべて無視し、形式幾何学上の分類基準を採用しています。この手法を選択したことにより、14の鍵を提案して、その鍵で旧石器時代の非表象的絵柄がすべて分類可能となるようにしました。このアプローチを進めるうちに、特徴のいくつかは無関係と見なさざるを得なくなりましたが、それでもなお3%の絵柄が分類不可能な「抽象的」絵柄として残りました。私達の分類は(現時点における)1つの可能な仮説と言うべきで、研究の進展につれ修正していく必要があるでしょう。ここで提案したタイプのほとんどが、ほぼ全地域で現出していることが観察できます。四角形は普遍的に思えますが、その素性のいくつかはスペインとフランスで異なっています。1つの四角形パターンにつき、四角形に特有な素性を5つ以上もつものは見られませんでした。さらに言えば、これら構成上の素性のいくつかは、独立したパターンとしてときおり現出しています。この結果、パターン表現と呼んでいるものが存在し、統合、重ね合わせ、並列という3つの構成タイプが識別できる、という結論に達しました。異なるタイプのパターン構成物(記号表現)が異なる内容(記号内容)を指示するものと解釈される、という仮説を立てました。旧石器時代の非表象的絵柄の内部構造に存在する、形式上の関係を表意的表記システムに見られる関係と比較しました。両者を比較するといっても、表記システムから何ら意味内容を借用する意図はなく、(内容とは独立して用いられる)ある種の形式上の手段を観察するのがねらいでした。こうした手段を用いたことにより、先史時代の人類には、抽象的パターン間に関係を確立するプロセスに(本来)存在する、2階抽象化の能力があった、という証拠も同時に得られたのです。

旧石器時代壁画の非表象的絵柄は、明らかに非線的であるという点で、表記システムとはかなり異なっています。その一方、大多数の「抽象的」パターンが動物の絵柄と相関してパネル中に現出しています。

動物のモチーフと抽象的パターンが、互いに影響を及ぼす2つの記号システムに属する、と見なせる証拠となる議論がいくつかあります。さらに、手形の陰画 (negative handprints) は多分、1つの複合記号システムの一部とも見なされるべきでしょう。

結果: タイポロジ、構成タイプ、相関規則

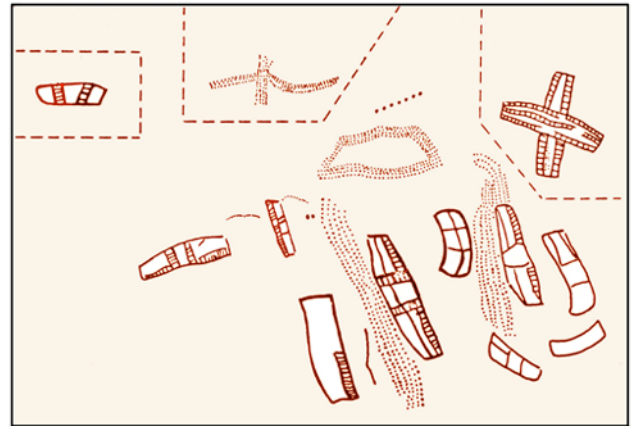


図1. エル・カステージョ洞窟(スペイン、カンタブリア地方) : 複数の非表象的モチーフの合成組立て(after H. Breuilにならって)

Design	Signifiers
	/ transverse ternary division / + / longitudinal binary division /
	/ transverse binary division / + / cross hatching /
	/ transverse ternary division / + / cross hatching /
	/ transverse ternary division / + / scaliform bands /
	/ transverse ternary division / + / scaliform bands / + / sharp-pointed excrescence /
	/ transverse ternary division / + / longitudinal binary division / + / scaliform bands / + / cross hatching /

図2. 類型(タイポロジ)上の鍵「四角形」に属する絵柄の解析。少数の描画記号表現から構成され、交換によって強調できる。

3. 統計的データ分析 – 因子分析とクラスタリングアルゴリズム (1988)

タスク： 表象的動物テーマを自動的に分類して、テーマ間相関を決定する

データ： 表象的テーマに還元した、パネル記述をもつデータベース

描画カテゴリ (モチーフ). ヨーロッパの旧石器時代洞窟芸術の表象コンポートは、14の主要モチーフに分別できます。(84の洞窟と岩穴シェルタにある、3300の表象的イメージを含む) データベースは、多重テーマ(2つから5つの異なるテーマ)をもつ416のパネルからなっており、統計学の観点から解析されています。数千の像がある種の規則にしたがって配置されているかどうかを論証するには、何らかの統計的処理が必要です。

私達のデータベースに存在する1027パネルのテーマ構成はあまり複雑でなく、その60%が単一テーマです(2つ以上のテーマを含むのは416パネルにすぎません)。複雑度が増すと、タイプの異なるパネルの出現頻度は急激に減少します。14の可能なテーマのうち、7つ以上を含んでいるパネルは存在しません。コミュニケーション理論ではよく知られていることですが、最も単純なものの出現頻度が最も高いものなのです。(Mandelbrot 1954).

要因解析 (Benze' cri 1984).

要因解析によると軸1の定義に最大の寄与をしているテーマは、片や雌ジカ、雄ジカおよび雄牛、他方ではマンモス、トナカイおよび熊です。この軸については、地域特性によって部分的に説明できるかもしれません。(スペイン)カンタブリア地方では雌ジカと雄ジカが多数いたのに対し、マンモスは主にフランス南西部に多くいたからです。第2の軸がマンモスを、稀なテーマ、すなわち、ライオン、魚および人工的カテゴリ「多様な」、から分離しています。馬、バイソンおよびアイベックスが多数の点の集まる、重心近くにあることは注目に値することです。

階層的クラスタ解析 (Jambu 1978).

要因解析はクラスタ解析で補完するのが通例です。クラスタ解析で得られるツリー構造が、要因解析にすでに登場していたクラスタリングをはっきりと確証します。この解析を通じて、群(またはクラス)がすぐさま確立されます。すなわち、レベル2で表現ツリーを「切断」することにより、5つの独立クラスが得られます:

- 馬、アイベックスおよびバイソン クラス1
- 人体の形象をしたもの クラス2
- マンモス、サイ、トナカイおよび熊 クラス3
- 魚、「多様な」およびライオン クラス4
- 雄ジカ、雌ジカおよび雄牛 クラス5

スパニングツリー アルゴリズム (Kruskal 1956).

Kruskalのアルゴリズムをテーマの共起関係表に適用することによっても、これら5クラスの一貫性は示されました。このアルゴリズムは完全グラフ、すなわち、テーマ間の可能なリンクすべてが考慮されているようなグラフ、から「最小コストのスパニングツリー」を抽出できます。

結果: 表象的モチーフの分類

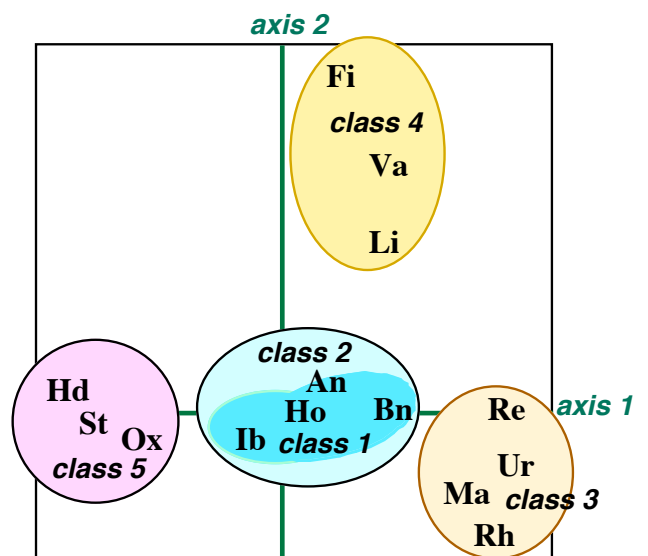


図3. 表象的モチーフ14個の共起関係表の要因解析

4. 自然言語処理 – 構文解析と文法推論(1995)

タスク: 表象的動物テーマの組み合わせをすべて決定する

データ: 表象的テーマに還元されるパネル

統計によって明らかになった構造には、下接する(隠れた)意味構造が反映されている、という仮説を立てます。このように仮定すると、5つのクラスを意味的クラスと見なすことができます。テーマの組み合わせ可能性についての制約は意味制約の結果です。私達の目的は、コーパス中で証明されているテーマの組み合わせを生成可能な、1組の規則を考案することです。この規則が旧石器時代洞窟芸術の形式文法、より正確に言えば、そのような文法のモデルを構成することになります。

ORグラフ文法。

第一歩として、最重要相関をすべて、自動的に見つけました。アルゴリズムは以下のとおりです。:

- コーパスは存在するタイプのパネルに還元する
- テーマの順序付けは、共起関係の頻度が減少する度合いによって行う
- 各テーマとリストで次にくるテーマとの相関を考慮するのは、頻度の降順で分類するためである
- 最も複雑なパネルでも6つのテーマしか含まれていないので、(最も出現頻度の高い5つの「規則主要部」に対応する--Ho, Bn, Ib, Ox--) 5つの主要規則のみを保持する
- 結果として得られるのは、-->が「OR」を意味するようなグラフである

この諸規則を適用する際、336タイプのパネルを生成できますが、コーパスには、そのうち106が存在しています。言いかえると、現存するパネルタイプの63%がこの規則によって説明されていることになります。さらに、生成されるパネルは426中

365(86%)に相当しますから、コーパス中で最も普通にあるものです。それゆえ、この非常に単純なアルゴリズムがシステムの優良な近似を与えてくれることとなります(前に述べたモデルは、6461の可能性のうち336の組み合わせのみを生産する、ということに注意してください)。このことから、集合が規則により支配されていることがはっきりと示されます。

文脈自由文法。

前に述べた、私達のORグラフ文法で表現される規則はシステム中核を抽象的かつ蓋然的方法で表示しますが、制約については何も述べていません。この理由から、(終端部を14の表象的モチーフとした場合、ダイアドを2つの終端部の合計、トライアドを1ダイアド + 1終端部、…、と定義する) DCG (Definite Clause Grammar、限定節文法) を記述しました。

クラスタ解析のおかげで、テーマの相関が、テーマの属しているクラスに強く依存していることに注目することになりました。それゆえ、これらのクラスは「構造上の制約」を公式化するために使用できます。5つの書き換え規則と少数の制約から成る、非常に単純なDCG文法では、制約をもたない文法が生産する組み合わせの、たった4分の1に相当する組み合わせしか生産されません。それでもやはり、洞窟で(その存在が)証明されている組み合わせの98%を生産できているのです。

結果: (Prologに組み込まれている、限定節文法の形式を用いる) 制約に基づく文脈自由文法と、自動推論されるOR グラフ文法

5. データベースからの知識発見(KDD) – ラフ集合近似(2001)

タスク: 表象芸術のパネルすべてを包括的に記述する

データ: 内在素性と外在素性をもつパネルの関係記述

今回は、ラフ集合近似値アルゴリズム (Pawlak 1982, 1991)を用いて、像を記述し、以下の属性から構成されるものとして、データを処理しました:身体部分(7価:全身(の動物)、頭部、無頭、無足、半身の頭部、後ろ半身、背中の線)、大きさ(5価)、向き(13価)、テクニック(14価)、アニメーション(5価)。でも、研究の現段階では、表象的モチーフと補い合ってシステム全体を構成する(と私達は信じていますが)、非表象的モチーフは考慮していません。

Pawlakのラフ集合理論は、私達の問題を解くのに非常に適していることがわかりました:つまり、ラフ集合理論によって、1組の未知の物体のうち、物体のもつ、どの属性が最大の識別力を示すかを決定できるのです。私達のデータにラフ集合アルゴリズムを適用すると、結果として、5つの属性を次のように順序付けて考慮することができました:

身体部分 > 向き > 大きさ > テクニック > アニメーション

この順序は、以下の最も出現頻度の高い動物(馬、バイソン および アイベックス)については不変であり、旧石器時代表象システムの定数(じょうすう)と見なすことができます。全く異種の動物を扱うには「身体部分」という属性を用います。私達のデータベースに最もよく出てくる2種の動物(馬とバイソン)に比較対象を限定するなら、馬というテーマが頭部で表現される方が、バイソンというテーマの場合より、はるかに多いことが観察できるはず(後者は無頭であることが非常に多く、あるいは背中の線だけで描かれていました)。

動物の向きが、2番目に重要な属性です。これはパネルの構成において重要な役割を果たしているようです。馬以外の動物種すべてに関して、ほとんどの像が左向きです。ただし馬は例外のようで、55,8%が右向きに描かれています。

パネルに、直接隣り合って描かれている動物の向きを考慮してみます。同一種に属する動物であれば、互いに向き合っていることは少なく、同一

方向を向いている可能性が最も高いことが観察できます。その一方、動物が2匹、前後に描かれているときは、前は馬で、バイソンが後ろ、ということが最も多く、その反対ではありません; バイソンと馬の向きが反対のときは、対面していることは少なく、交差していることが最も多いのです。しかしながら、私達のテストベンチ データベースはまだ小さすぎて、こうした観察に確証を与えるまでには至っていません。

結果: フランコ=カンタブリア北部旧石器時代芸術の、非常に有意義なテストベンチ情報データベースに対して、ラフ集合計算法を備えたKDD手法を適用し、少数の重要な記述素性を識別できた

6. 予測

タスク: (地理的、通時的) 下位構造を識別する

データ: (表象的絵柄と非表象的絵柄を関連づけて) データを更に修正する

本研究の次なるステップは、個々の像を可能な限り多くの属性によって記述する、巨大データベースを構築することです。このタスクは進行中ですが、2つの理由から困難だといえます。まず、必要な情報のいくつかは文献に無く、収集しなければならないこと。第2には、将来の研究において、データマイニング技術と機械学習(マシンラーニング)アルゴリズムをもっと活用して、注意深く芸術の領域に適用する必要があることです。

現在はこの代わりに、旧石器時代岩面芸術を、アフリカおよびオーストラリアの民族的データと統計的に比較する研究を進めているところです。

7. 結論

意味解釈の可能性を客観的方法で獲得することは、非常に困難です。別の意味体系に対して(例えばシャーマニズム、またはトーテム崇拝タイプの信仰)、文書化を伴わない類推を行うことは避けるべきだ、ということは十分承知しています。(色、

テクニック、大きさなど) できるだけ多くの属性を考慮することにより、複数の像の間に(類似性や対立などの)複合関係が存在していたという、人の知能に関する3万年以上昔の証拠を提出することができるのですが、複合関係がもつ意味内容は、多分ずっと明らかにされないまま残ることになるでしょう。

重要な問題をいくつか、ここで要約しておきます:

- 1) ヨーロッパ旧石器時代洞窟芸術における表象的モチーフは数少ない。馬、バイソンおよびアイベックスのようないくつかのモチーフは、真のライトモチーフである(それらを合わせると、画像の60%以上を代表することになる)
- 2) 生産された表象テーマの相関は、可能な組み合わせ数に比べるとはるかに小数である
- 3) これらの特徴は少数規則から成る形式文法により説明可能である。非常に単純で一貫したモデルが派生され、これで旧石器時代洞窟芸術作品の98%を説明できる

少数の一貫した繰り返し規則によってコーパス全体が説明できる、という事実には、旧石器時代洞窟芸術を代表とする非常に広大な地域で、何千年にもわたって、同一の複合規則が使用されてきたこと、が含意されています。

私達の戦略は、逐次近似法によって一步一步前進することでした。第1次近似値として、旧石器時代洞窟芸術には共通構造が存在したのかという、重要で議論を呼ぶ問題に答えることが目標でした。

ラフ集合理論によって、最も識別力のある属性を自動決定し、有意水準リストを進歩的に修正する、アルゴリズムが与えられます。この手法を適用すると、モチーフの向きや相対位置のような、位相相関属性の重要性を指摘できるようになります。2つの最重要モチーフである馬とバイソンは、これらの基準に関して異なる特性を示しますが、

これが構成素性と解釈できることは注目に値する、興味深いことです。私達は表象統語論の基礎を発見したばかりのところにいるのです。

References

- [Leroi-Gourhan 1965] Leroi-Gourhan A., *Préhistoire de l'art Occidental*, Editions Mazenod, Paris, 1965 (English translation: *Treasures of prehistoric art*. Abram, New York)
- [ソヴェー 1977] G. ソヴェー、S. ソヴェー、A. ヴロダルチック、「先史時代の記号学式論」一人類最初の猫記符号の理論のために」フランス先史時代学会紀要、1977年、74巻、研究論文2号、パリ「エピステーメ」月刊誌、1977年、8月号、朝日出版社、東京
- [Sauvet 1977] Sauvet, G., S. Sauvet & A. Wlodarczyk, *Essai de sémiologie préhistorique (pour une théorie des premiers signes graphiques de l'homme)*. *Bulletin de la Société Préhistorique Française* 74: 545-58, 1977.
- [Sauvet 1988] Sauvet, G., *La communication graphique paléolithique (de l'analyse quantitative d'un corpus de données à son interprétation sémiologique)*. *L'Anthropologie* 92: 3-16, 1988.
- [Sauvet 1992] Sauvet G. & Wlodarczyk A., "Structural Interpretation of Statistical Data from European Palaeolithic Cave Art", in "Ancient Images, Ancient Thought: the Archaeology of Ideology" Proceedings of the 23rd Chacmool Conference, Calgary, Canada (Nov. 1990), S. Goldsmith, S. Garvie, D. Selin, J. Smith Editors, The University of Calgary (Canada), 541 p. (223-234), 1992
- [Sauvet 1995] Sauvet, G. and A. Wlodarczyk, *Eléments d'une grammaire formelle de l'art pariétal paléolithique*. *L'Anthropologie* 99: 193-211, 1995.
- [Sauvet 1999] Sauvet G. & Wlodarczyk A., "Towards a Formal Grammar of the European Palaeolithic Cave Art", "Proceedings of the 4th World Symposium on Palaeolithic Cave Art" (CD-ROM), Pinerolo, Italy, 1999 (to be reprinted in "Rock Art Research", AURA, Australia).
- [Sauvet 2001] Sauvet G. & Wlodarczyk A., "L'art pariétal, miroir des sociétés paléolithiques". *Zephyrus*, n° 53-54, p. 215-238.