

「分かる」を支援するAIとは？

溝口理一郎

北陸先端科学技術大学院大学

サービスサイエンス研究センター



Agenda

- AI
 - 歴史
 - コンピュータ処理と意味
- 分かるとは？
- 「分かる」を支援するAI
- 機能分解による知識記述
- 磁石材料設計知識への応用



AIの主な研究課題

■ 意味理解

- 知識表現
- 自然言語理解, 対話システム
- 画像理解, パターン理解

■ 推論

- 述語論理, 演繹推論, 帰納推論

■ 学習

- Neural Network, 多くの学習技術

■ 知識ベース

- 知識工学, オントロジー工学
- Semantic Web

■ 応用システム

- ゲーム, 自動運転, 自動XYZシステム



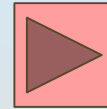
第一次のブーム

- コンピュータの誕生の10年後(1960年代)
- **GPS: General Problem Solver**
 - 状態空間
 - 状態変化Operator
 - 初期状態をゴール状態に変換する
- **前提:** どんな問題でも, その解決過程は原理的には状態遷移として表現できる
- **困難さ:** 計算時間が指数関数的



第二次のブーム

■ エキスパートシステム(1980年代)



- 専門家の経験則を知識ベース化
- 説明能力のある推論エンジン
- 実問題で人間の専門家と同等以上の性能発揮
- **実例**: MYCIN(感染症患者への抗生物質投与)
石油掘削支援, 電気機関車の故障診断, 生産工程計画など

■ 前提: Knowledge is Power(Ed Feigenbaum)

知識工学の誕生→オントロジー工学への発展

■ 困難さ: 知識獲得(知識ベースの構築)

第三次のブーム

- Deep learning (特徴抽出学習) 
 - 単に多くの中間層があるNeural Networkではない
 - 将棋や囲碁ソフトの強さの増強に貢献
 - データ科学の本質的技術
- 前提: データは知識の宝庫
大量のデータがある, 強力な計算パワーがある
- 利点: 最も困難な**特徴抽出**の自動化を実現
- 困難さ: Black box → 説明能力の欠如
大量のデータが無いと動かない

コンピュータ処理

■ アルゴリズム（手続き指向）

- 数値計算
- 記号処理

■ 探索（状態指向；状態＝記号）



- 定められた範囲の探索空間を探し回る
- 推論エンジン



記号処理における意味



知識表現

[もの]	[関係]
block A	on(X, Y)
block B	above(X, Y)
block C	clear(X)
Table	holding(X)
hand	handEmpty

[操作]

つかむ, 置く

if 操作前状態

then 操作後状態

つかむ=

if on(X,Y), clear(X), handEmpty

then clear(Y), holding(X)

block A, Table, hand A
on, clear, holding...
の意味は不明！

操作が合理的である範囲という制約で
意味が暗黙的に規定されているだけ

データ構造 + アルゴリズム

知識表現 + 推論

論理(推論) と 意味

$$\begin{array}{c} \blacksquare \quad P \rightarrow Q \\ P \\ \hline Q \end{array}$$

$p(X,Y):-f(X,Y).$
 $g(X,Y):-f(X,Z), p(Z,Y).$
 $f(a,c).$
 $f(c,b).$
 $\Rightarrow g(a,b).$

- $p(X)$
 - $human(X)$
 - $tall(X)$
 - $hungry(X)$
 - $walk(X)$

$parent(X,Y):-father(X,Y).$
 $grand_father(X,Y):-father(X,Z), parent(Z,Y).$
 $father(john,tom).$
 $father(tom,bob).$
 $\Rightarrow grand_father(john,bob).$

内容/意味: オントロジー



AI技術 二つの側面

- 性能指向(工学的)
 - 新しい(有用な)ものができれば良い
- 理解指向(科学的)
 - 「分かる」ことが大切
- 知識 + 推論
アーキテクチャ(自意識)



分かるとは何か？

- 将棋でプロに勝ったからと言って将棋が分かったことにはならない
 - Q:なぜ, ここでこの手を打ったのですか？
 - A:評価値が一番高かったから
 - Q:なぜ, 評価値が高かったのですか？
 - A:??
- 説明性の欠如
- Blackboxと言われる
(既にアメリカでは「説明できるAI」プロジェクトが立ち上がっている)



説明とは何か？

概念1 → 概念2 → 概念3 → . . . → 概念N

- 人が理解できる概念が使われていること
- 概念と自分が持つ知識との整合性がとれること
- 論理的な繋がりが妥当であること
- ある事象の原因が分かること
 - 因果的な繋がりが妥当であること
- **因果**とは？ 原因とは？

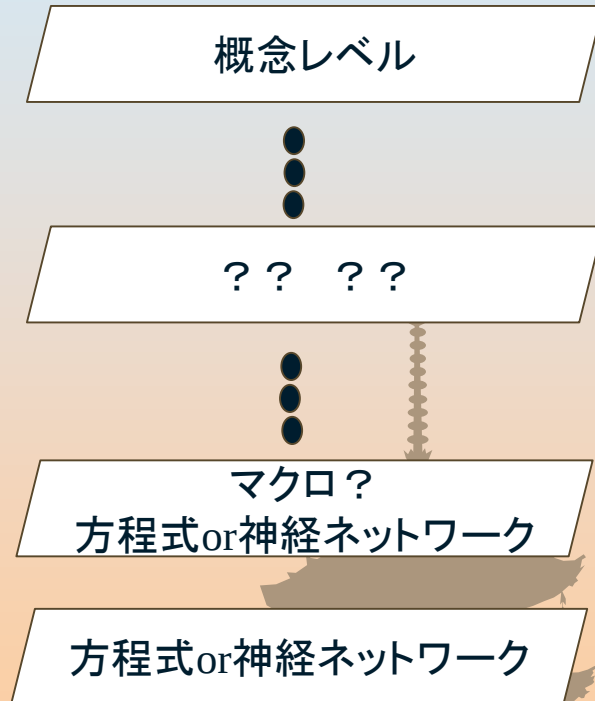


Functional square of causation

	direct	indirect
positive	Achieve	Allow
negative	Prevent	Disallow

学習を使った記述子の獲得

- 第一原理(方程式)は, 説明対象との距離が遠いので説明力が不十分
- 機械学習による記述子の学習?
 - Deep learningの真骨頂
 - Autoencodingによる特徴の自動抽出
 - 大量のデータがあることが前提
 - 本質的に数値計算
 - 説明能力は無い
 - 相関と因果は別物なので
- 事前に用意した記述子の中から学習によって選択



概念を根本から議論する

- 存在論
- 概念は所詮，存在するものに「**関する**」何か



オントロジーとは？

- 哲学: 存在に関する体系的な理論(存在論)
- 人工知能: An explicit specification of conceptualization
- 人間が対象世界をどのように見ているかという
根元的な問題意識を持って物事をその成り立ちから解き明
かし, それをコンピュータと人間が理解を共有できるように書
き記したもの
- オントロジーは,
客観的な存在の諸相を明らかにする立場から基盤となる
概念体系を整備して, **意味を扱う理論と方法論を提供する**
- 概念研究の親玉 ☺



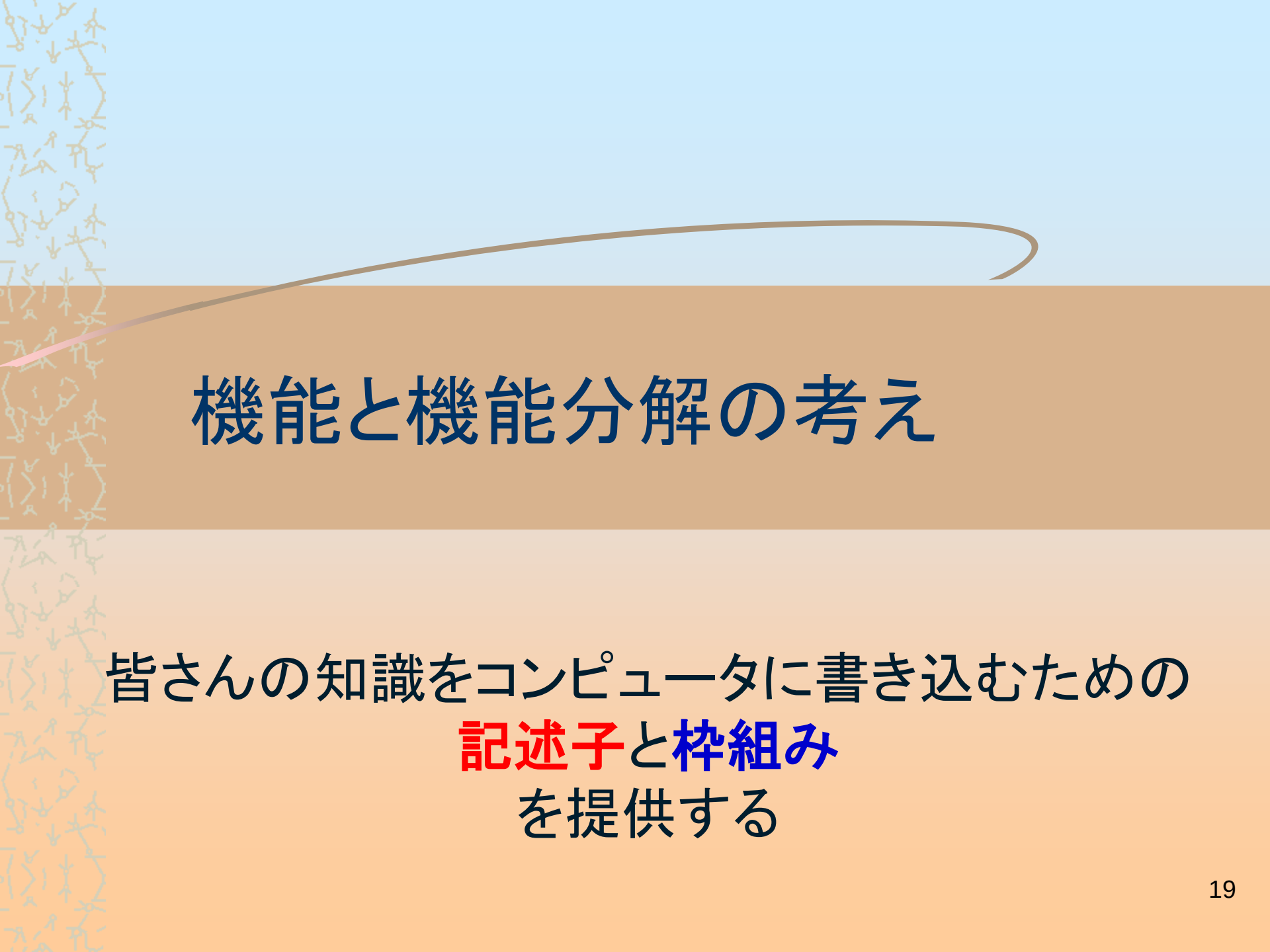
YAMATOの上位の 概念構造



オントロジー工学

- 「分かる」ことを中心に据えた本格的AI理論と技術
- 「分かる」の本質は「説明」にある
- 説明の本質は**概念**と**因果関係**にある
- オントロジー工学は**概念**を，存在から説き起こして，問題に横たわっている
 - 本質的な**概念構造**を明らかにする
 - 同時に，思考に**有用な「枠」**を与える





機能と機能分解の考え

皆さんの知識をコンピュータに書き込むための
記述子と**枠組み**
を提供する

中心となる概念

従来の知識ベース・知識工学

■ デバイスオントロジー

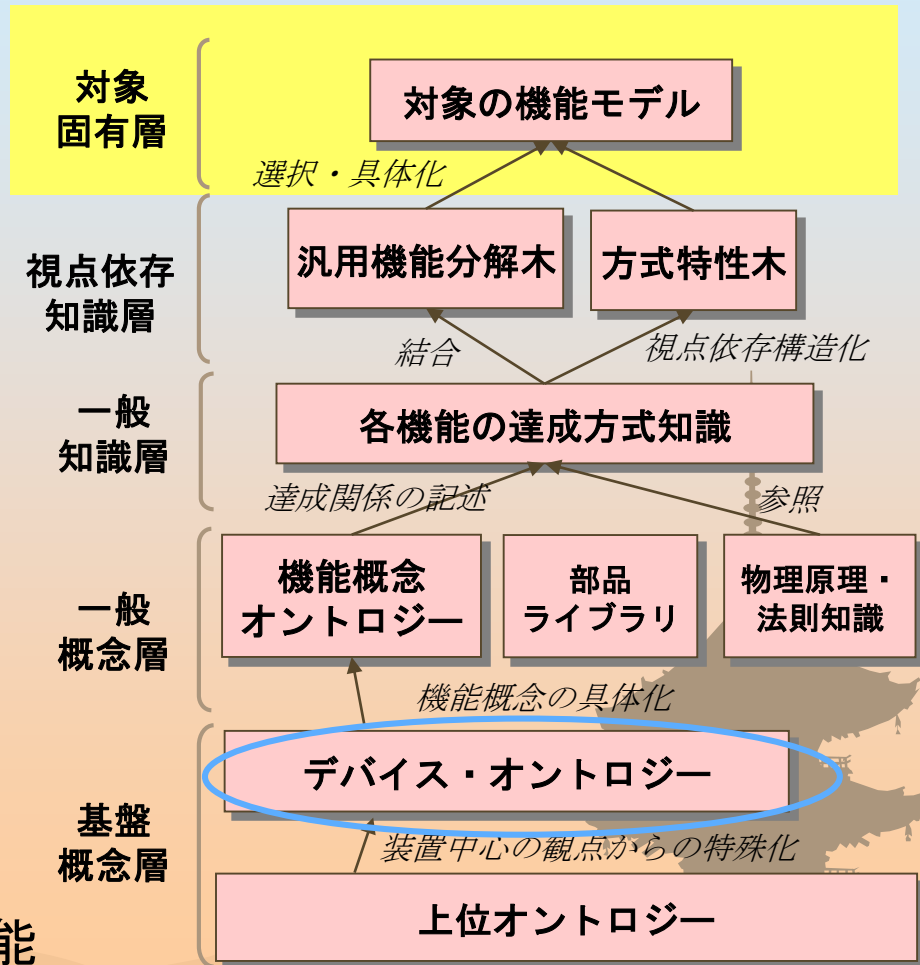
- Agent, 対象物
- デバイス(agent), 導管, 流れるもの(対象物), 媒体

■ 機能モデリング

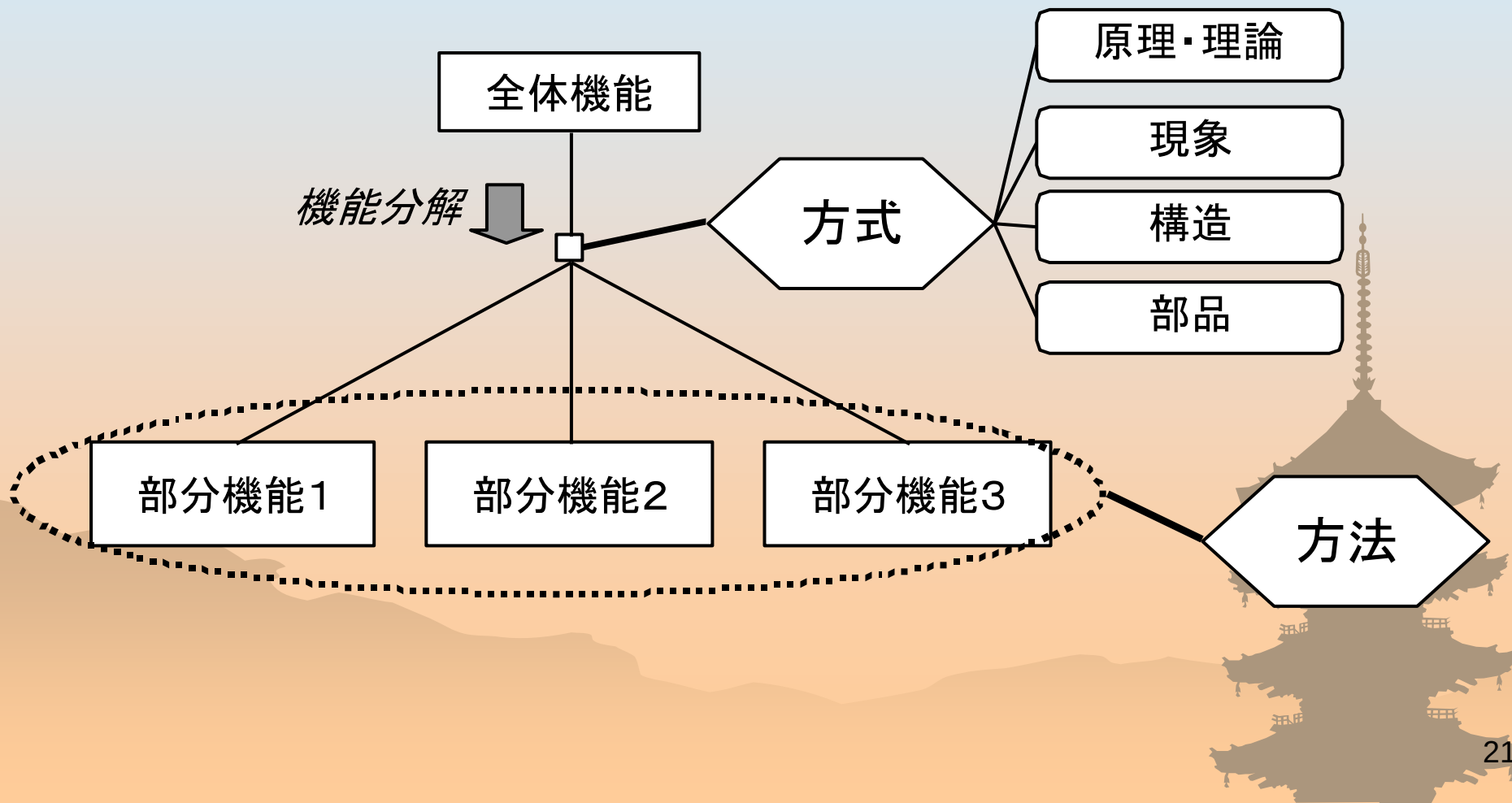
- 振る舞いと機能
- メタ機能と機能タイプ
- 機能分解
 - ・ 方式と方法

■ 知識のモデリング

- *Is-a* 階層: 機能と方式
- *Part-of* 階層(*is-achieve-by*): 機能



機能分解：方式と方法



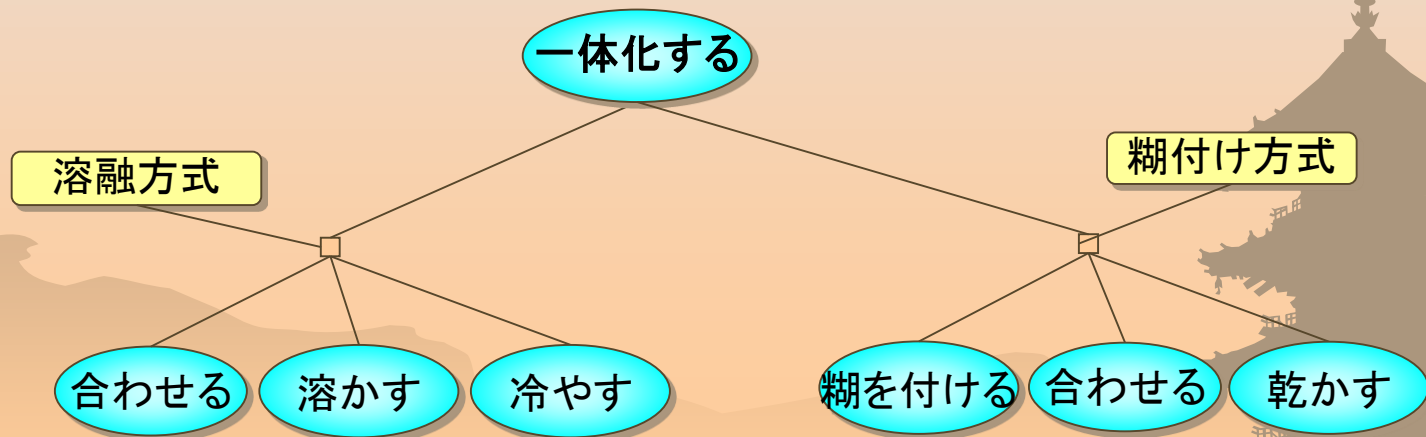
例

■ 溶接する

- What to achieve = 2片の金属を「一体化」する
- How to achieve = 溶融方式

■ 糊付けする

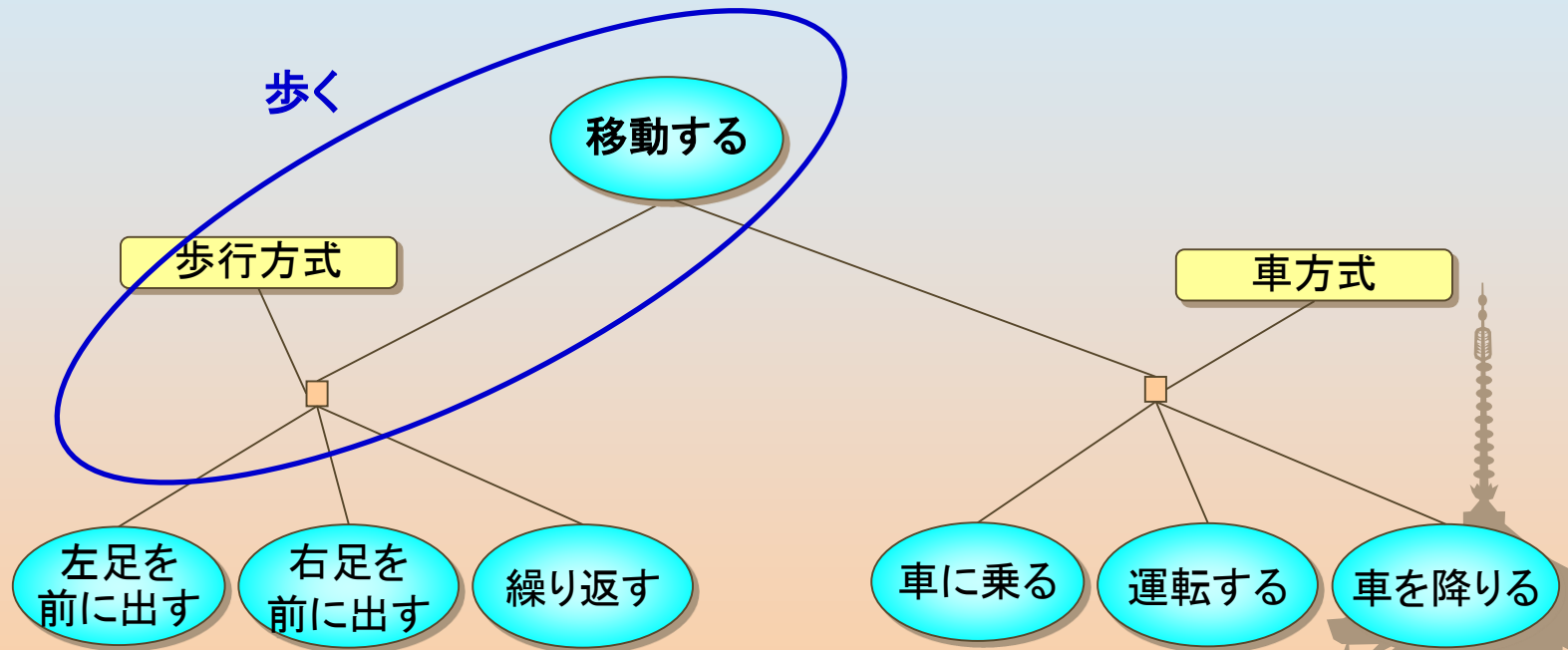
- What to achieve = 2つのものを「一体化」する
- How to achieve = 糊付け方式



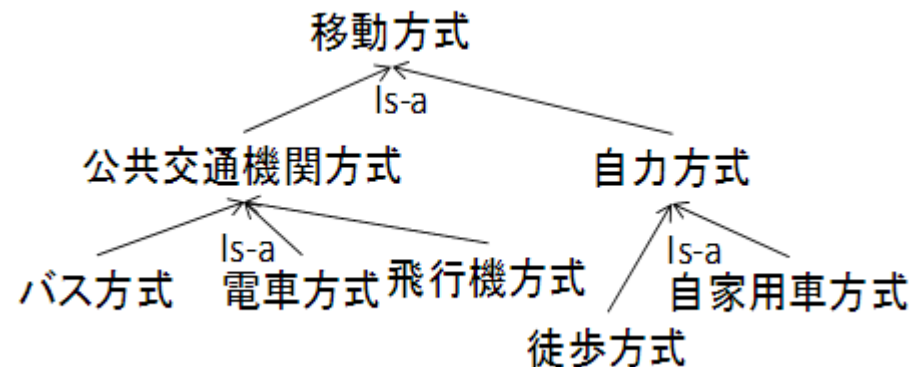
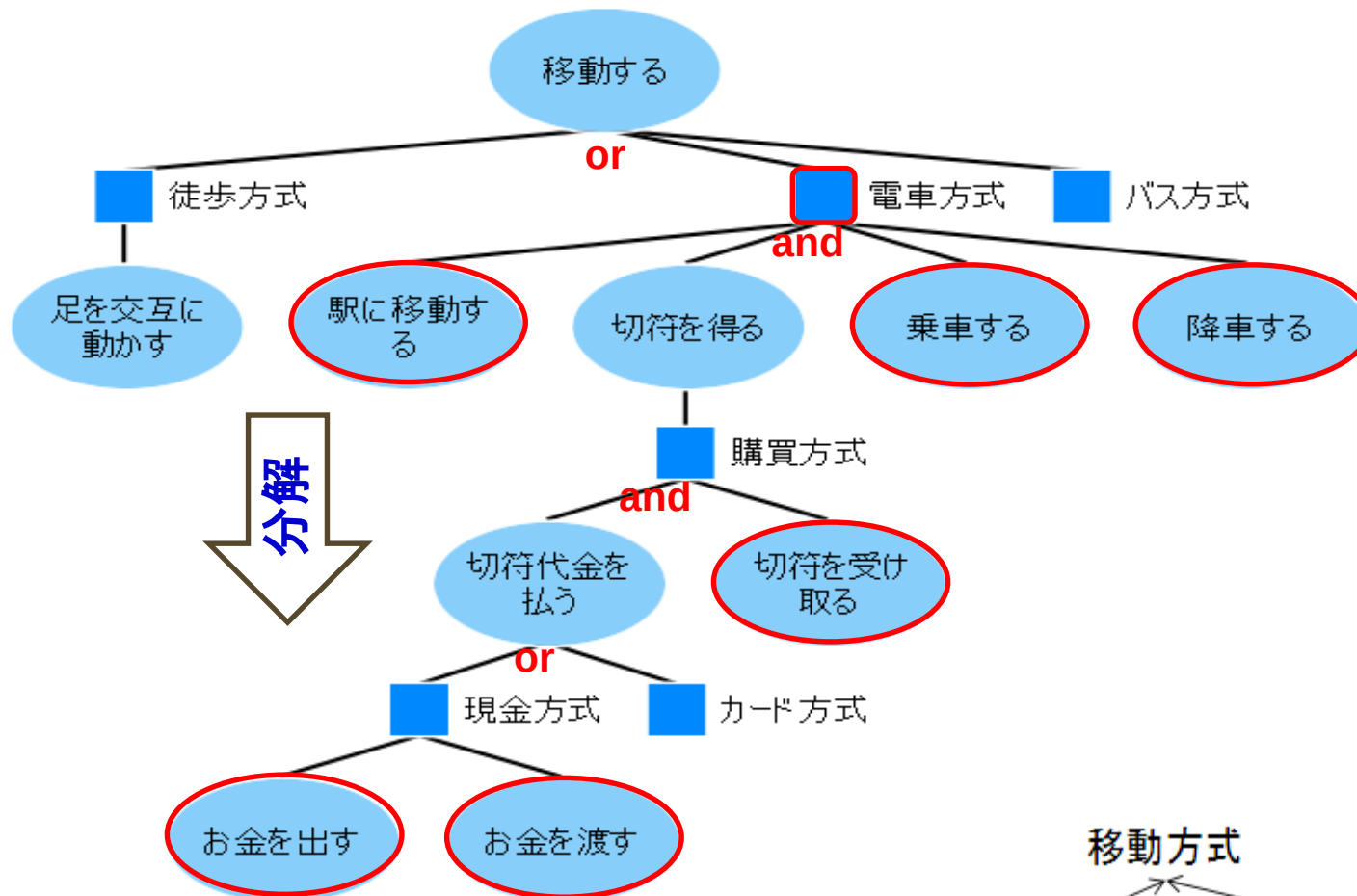
行為の分解

- 「歩く」は行為ではない！？
- 「歩く」は移動のための手段(方式)である
- 「**移動する**」は「居場所を変化させる」という行為であり、その実現手段(方式)を特定しない、状態変化だけを概念化したものである
- 移動する行為の実現方式は「歩く」以外に、**走る、車を運転する、電車に乗る、飛行機に乗る、**などたくさんある

移動行為の分解



「移動する」の行為分解木



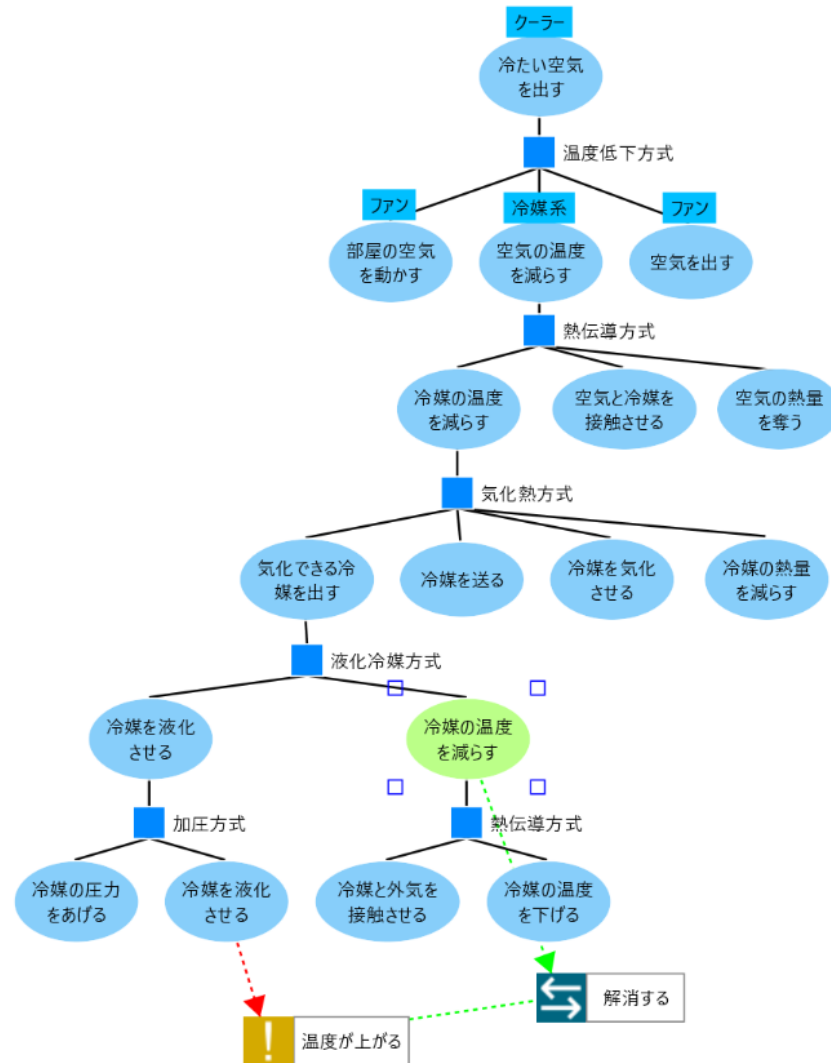
OntoGearCore

ファイル(F) 編集(E) 表示(V) 操作(O) ツール(T) ヘルプ(H)

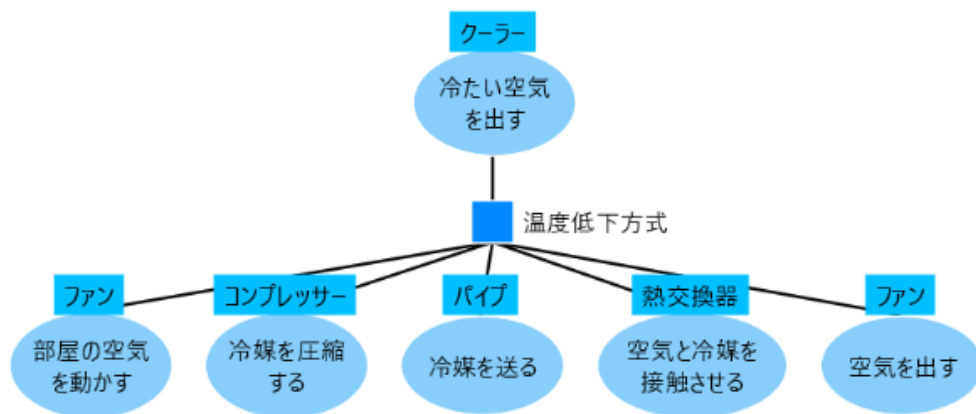
まずい例 良い例

機能プロパティ

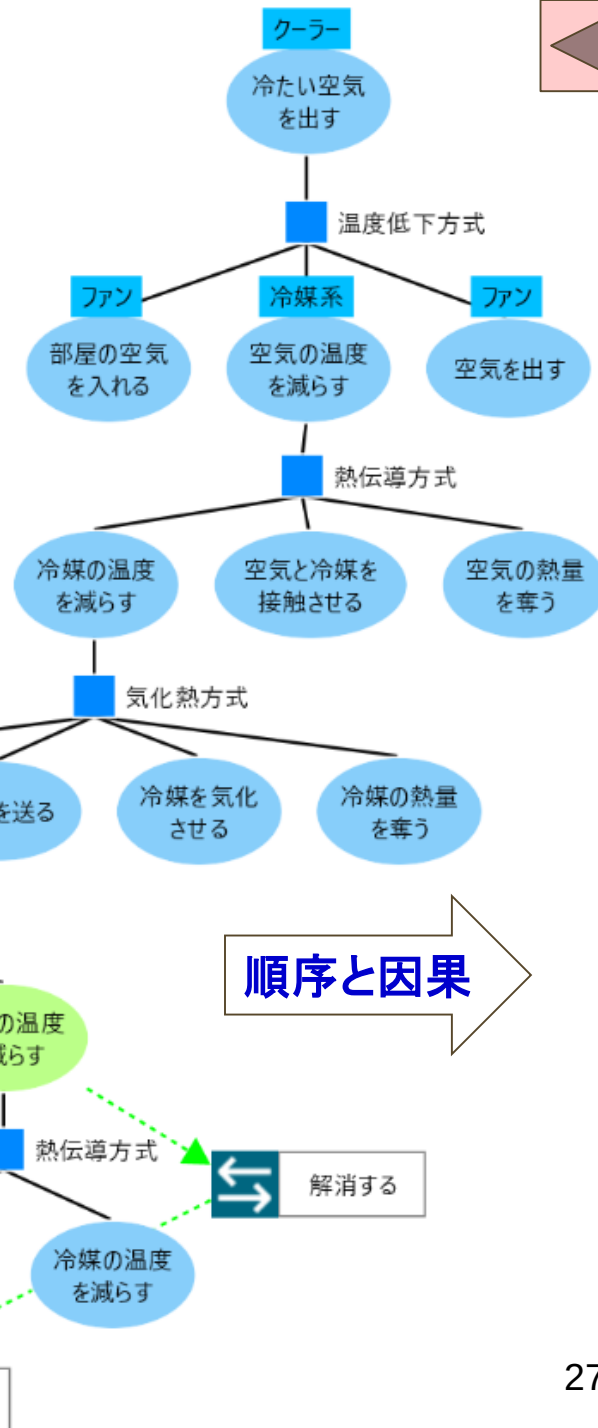
主体	☐		▼
機能表記	冷媒の温度を減らす		
機能	減らす		
対象物			+ 削除
	冷媒の温度	▼	を ▼
機能種別	プロセス維持		
作成日	2018/05/06		
属性			+ 削除
条件			+ 削除
制約			+ 削除



クーラーの機能分解木



細粒度



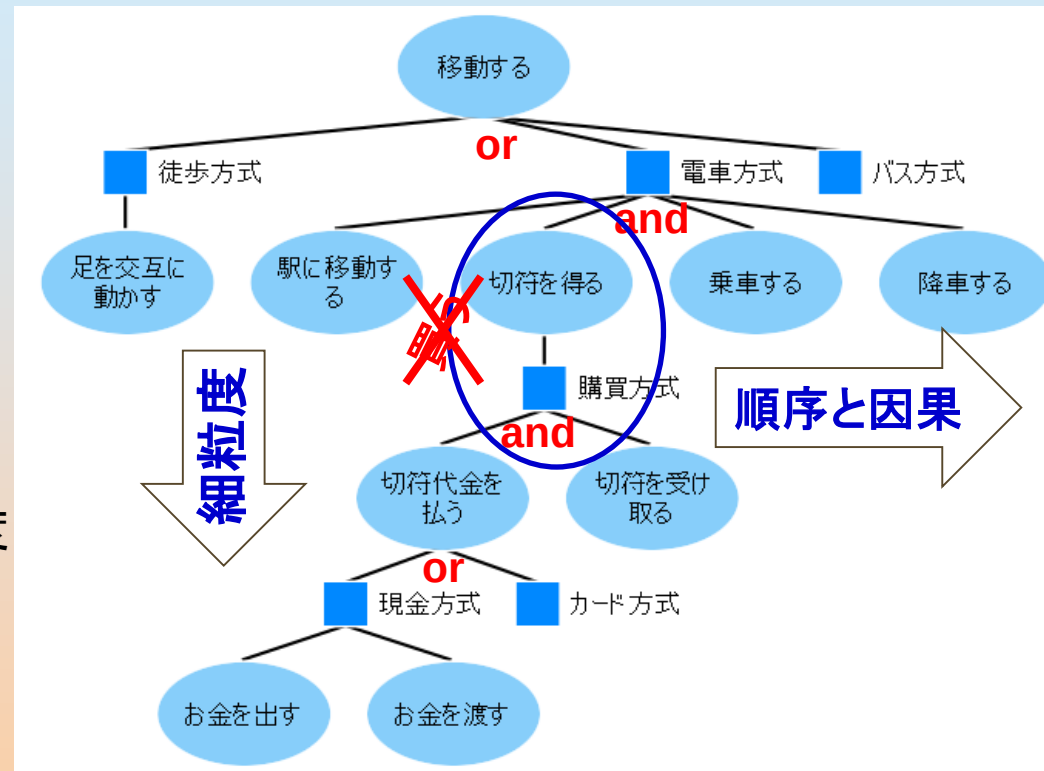
順序と因果

機能(行為)分解の肝と意義

肝



- 現象は**状態**とその変化で捉える
- 「やり方」は**方式**に押し込める
- 上のノードは**ゴール**と読む
- 下のノードは行為と読む
- 因果や時間は左から右に流れる
- 上下は**粒度**の大小: 説明の詳細度
- 切符を「**買う**」と言ってはダメ
 - 得るという



意義

- 隠れていた(中間)ゴールを浮き彫りにする
- 同じゴールに複数の方式を紐付けする
 - 方式の置き換えで新しい手法を作れる

得る(人が行っていること)

得る

■ (方式)

ー 選択する

- 無評価方式

- ー 候補からランダムに選ぶ

- 評価方式

- 1. 評価基準を選択する or 作成する(再帰) and 評価する

- 2. 評価値を参照して選ぶ

ー 作成する:(部品=操作可能な物)

- 1. 部品を選択する(再帰)

- 2. 部品を変形 or 分解する

- 3. 部品を組み合わせる(一つだけの場合は選択)



磁石材料関連知識への適用

- NIMSの 真鍋さん と 木野さん との共同研究



1. はじめに

ネオジム磁石は Fe, Nd, B を主成分とする合金から製造され、 $\text{Nd}_2\text{Fe}_{14}\text{B}$ 立方晶化合物を主相とする磁石の総称である。ネオジム磁石はこの主相の結晶磁化容易軸（c 軸）が無配向の等方性磁石と、一方向に配向した異方性焼結磁石の 2 種類に大別される。等方性磁石は液体急冷法で製造され、図 1 に示すように結晶磁化容易軸が等方的であるために、回転磁化により初磁化曲線の立ち上がりが悪く、減磁曲線の角形性が悪い。結晶粒間の交換結合が弱いと残留磁化は飽和磁化の 1/2 となり、高いエネルギー積は得られない。しかし、結晶粒が磁壁のピンングサイトとして作用するので、容易に高保磁力が実現できる。一方で、異方性磁石は図 1 に示されるように、結晶磁化容易軸が一方向に強く配向しているために磁壁移動による初磁化曲線の立ち上がりが良く、減磁曲線の角形性も良好になる。異方性磁石の代表的なものは焼結磁石であり、粒内に磁壁のピンングサイトがないために保磁力は低い。 Nd-Fe-B の 3 元素だけで製造される焼結磁石の保磁力は 10 kOe 程度である。別種の異方性磁石としては、水素不均化脱離再結合法(Hydrogen disproportionation desorption recombination, HDDR)という $\text{Nd}_2\text{Fe}_{14}\text{B}$ を主相とする粉体を水素化し NdH_2 , Fe_2B , Fe の 3 相ナノ組織を形成し、それを再度脱水素化することにより $\text{Nd}_2\text{Fe}_{14}\text{B}$ を再結合させて、200 - 300 nm の超微細結晶粒が強く配向した異方性磁粉がある。

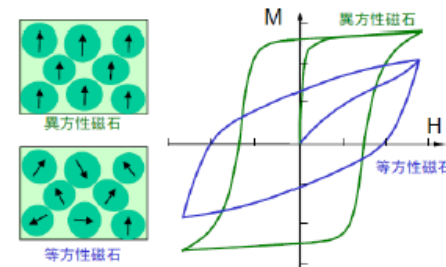
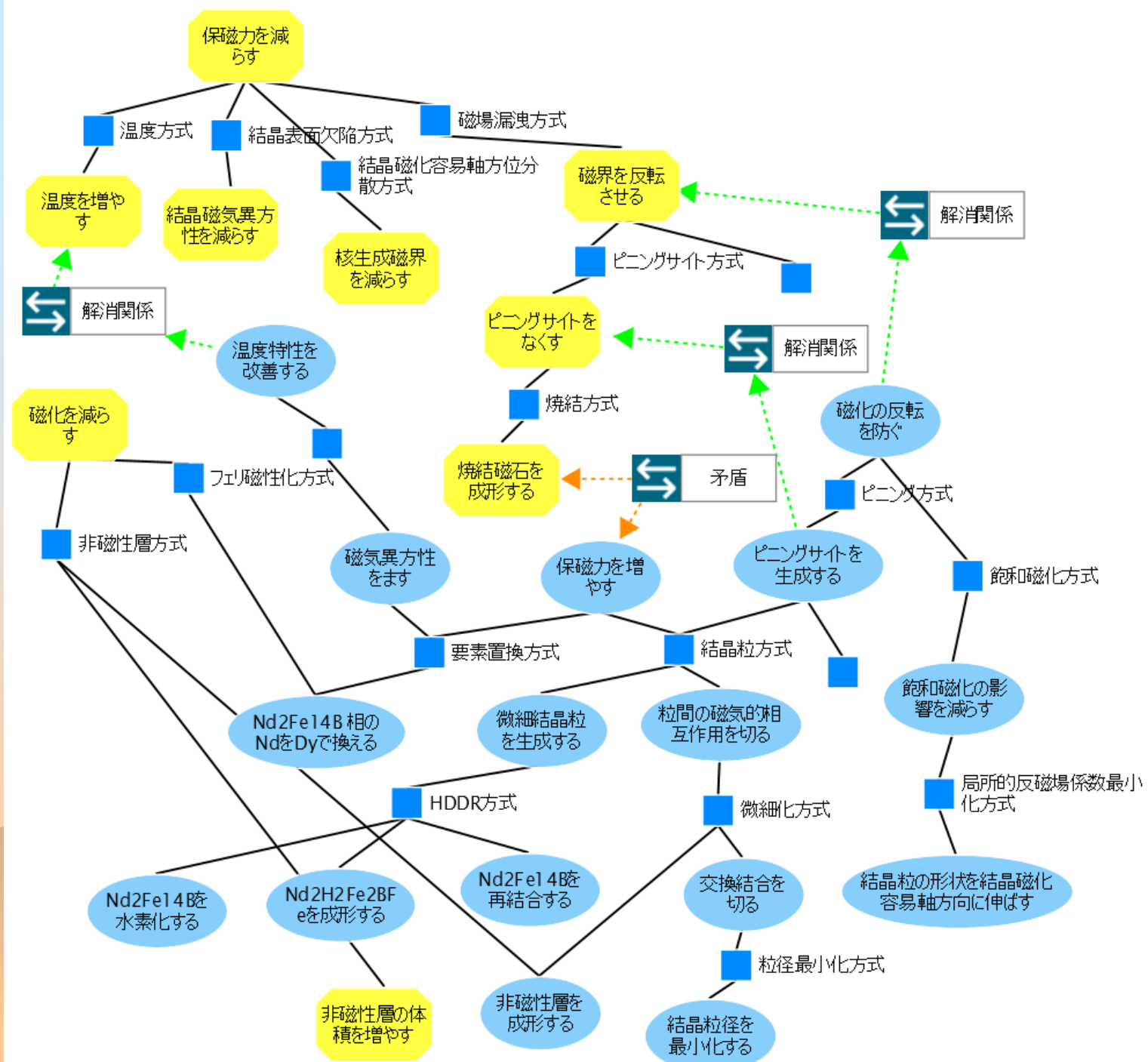
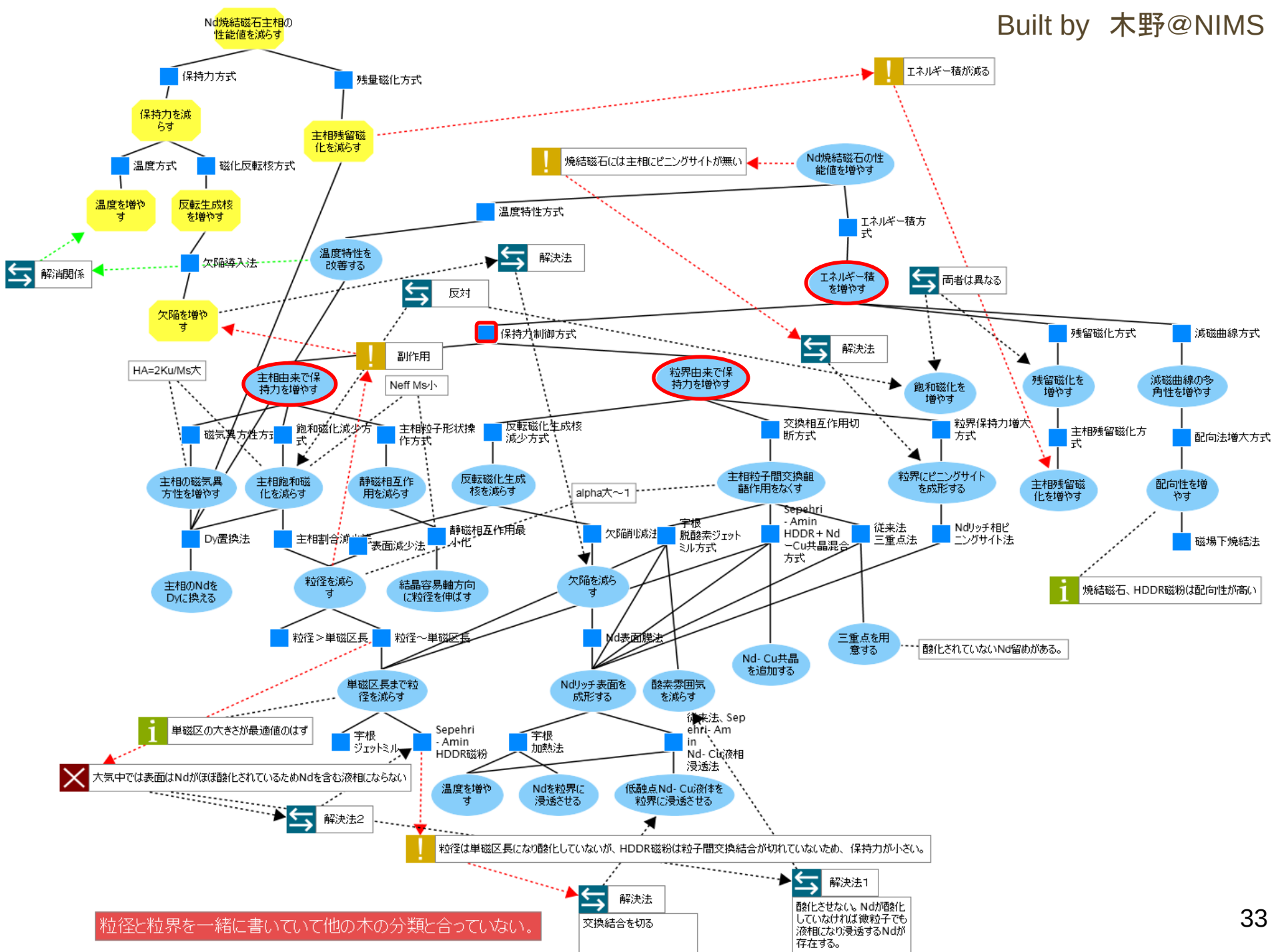


図 1 等方性磁石と異方性磁石の微細組織の模式図と典型的な磁化曲線

これらのネオジム磁石の中でとりわけ重要なのが焼結磁石で、その配向性の高さから減磁曲線の角形性が良好で、理論限界の 90%を超える 55 MGOe もの最大エネルギー積の高性能磁石が工業的に生産されている。現在のところ、焼結磁石だけがハイブリッド車・電気自動車(HV/EV)用の駆動モータと風力発電機の小型化・高性能化に必要な高特性を発揮できる。最大エネルギー積 50 MGOe 級の焼結磁石は極微量 Cu と Al を添加した Nd-Fe-B 三元合金を基本としており、その保磁力は 12 kOe 程度である。磁石の保磁力は温度が上がると低下し、室温で 12 kOe の保磁力もモータの動作温度の 200°C ではわずか 2 kOe 程度にまで低下し、駆動モータの使用環境で減磁してしまう。

この問題を解決するために開発された磁石が Dy 含有焼結磁石であり、 $\text{Nd}_2\text{Fe}_{14}\text{B}$ 相の Nd の一部を Dy で置換することにより主相の結晶磁気異方性を高め、それにより保磁力を高めている。約 30%の Nd を Dy で置換した焼結磁石の保磁力は 30 kOe 程度になる。ところが Dy 置換した $(\text{Nd,Dy})_2\text{Fe}_{14}\text{B}$ 相は Nd と Dy のスピンの反平行結合がフェリ磁性的となり、磁化が下がってしまう。そのために最大エネルギー積は 30 MGOe 程度にまで下がってしまう。現在、この 30 MGOe 級の Dy 含有高保磁力磁石が HV/EV、風力発電用磁石として使用されている。





保持力の低下要因

By 真鍋さん

■ 省略



キューリー温度予測問題

■ 3つの手法の分解木による比較

1. Schefflerたちの方法
2. Dam先生の方法
3. 木野さんの方法



Ghringhelliらの手法

目的: 82 binary metalのZinc-blede (or wurtzite) vs rocksalt
構造の全energy差を線形モデルで回帰する。

14

個の基本原子記述子を用意する。

関数適用

 10^4

個の記述子

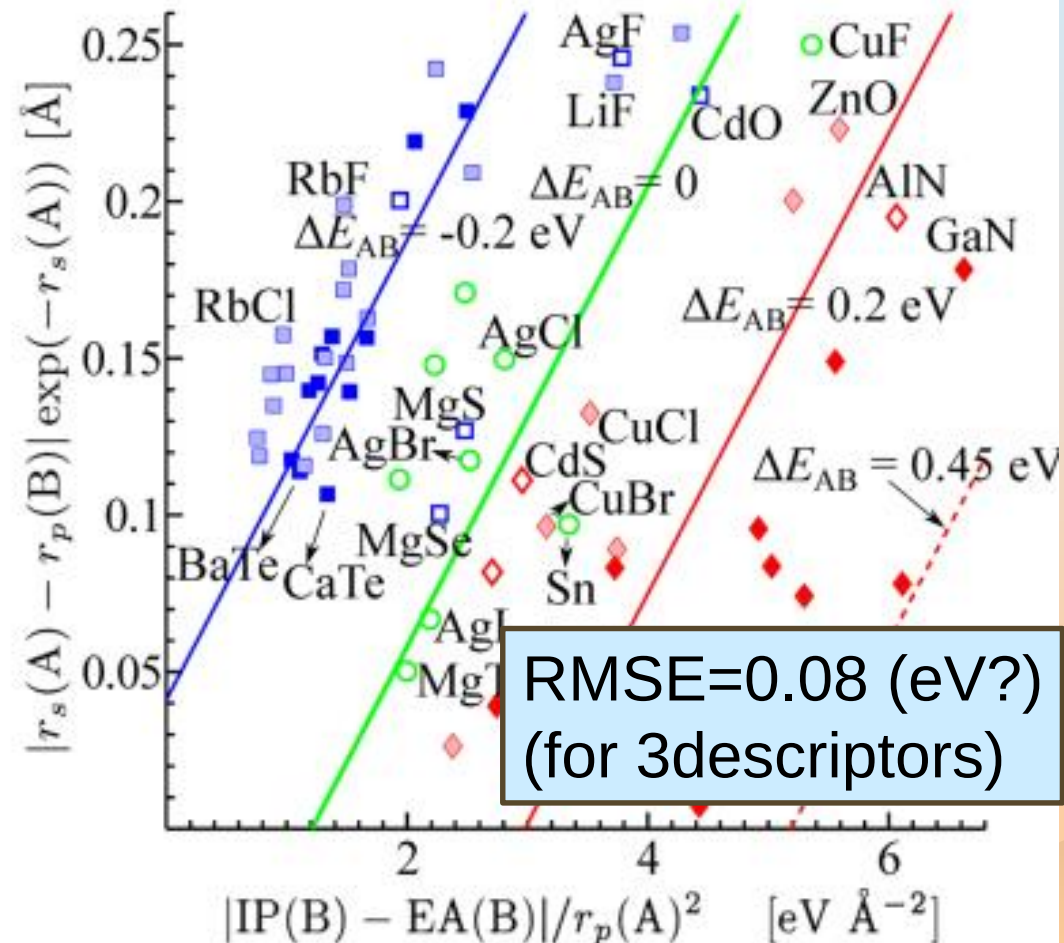
LASSO

25-30

個の記述子

全探索

数個の記述子を得る。



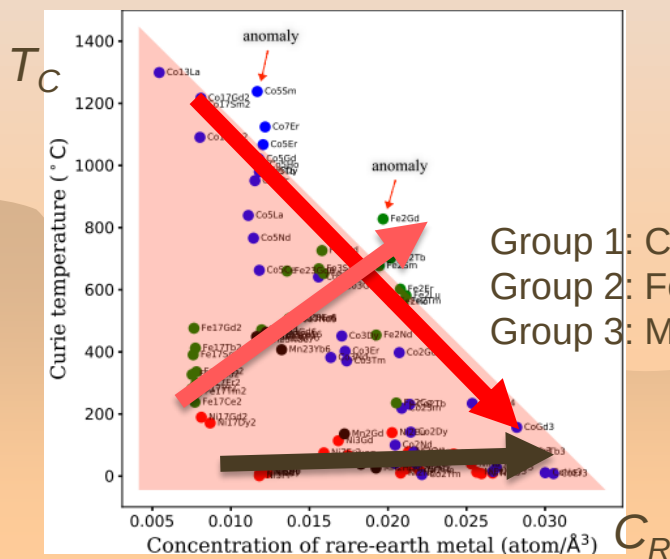
Regression based clustering

(Dam先生の方法)

- 108 binary alloyのCurie温度を区分線形回帰す。
- 27原子、構造記述子を用いて全探索を行う。
- 3-4記述子で線形回帰による分類が可能になる

それぞれの区分線形モデルの重要な記述子の係数

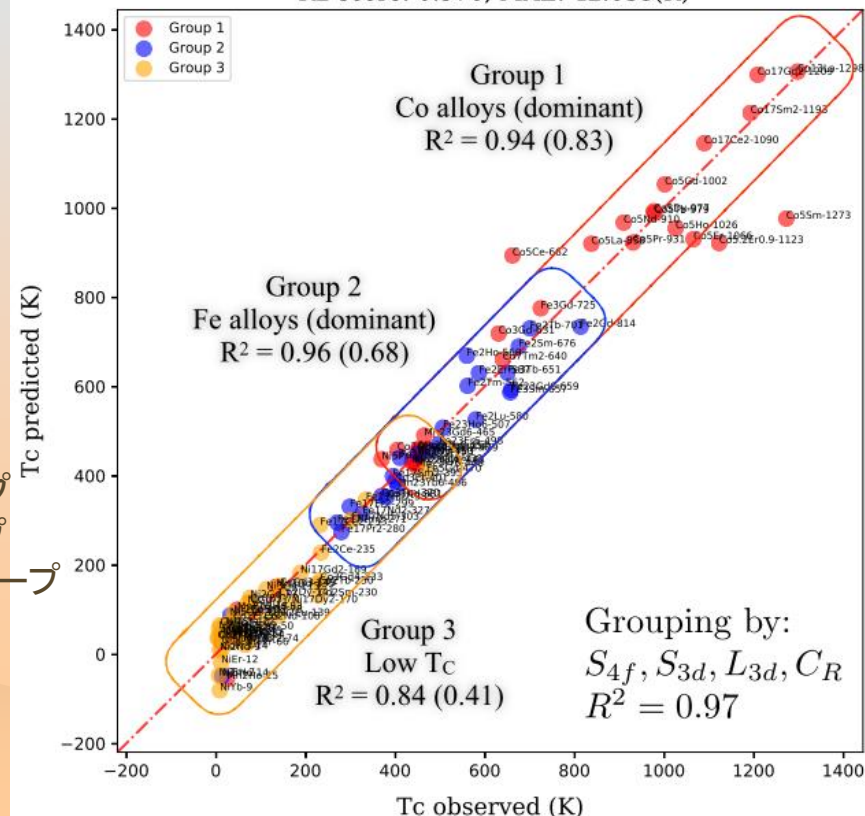
	S_{4f}	S_{3d}	L_{3d}	C_R	Intercep
Group 1	46.8	105.9	522.8	-66661	-1502
Group 2	74.1	664.3	163.4	23610	-1621
Group 3	55.2	131.2	-124.83	-4875	274



Group 1: Coを主としたグループ
 Group 2: Feを主としたグループ
 Group 3: Mn, Niを主としたグループ

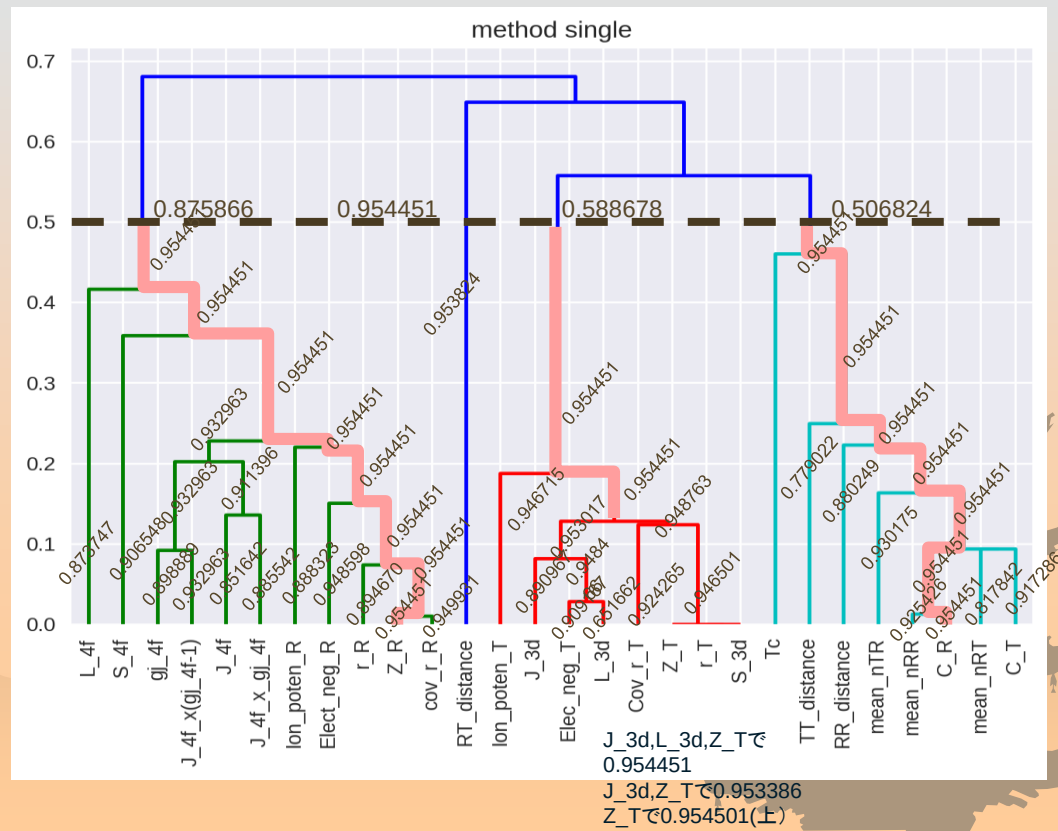
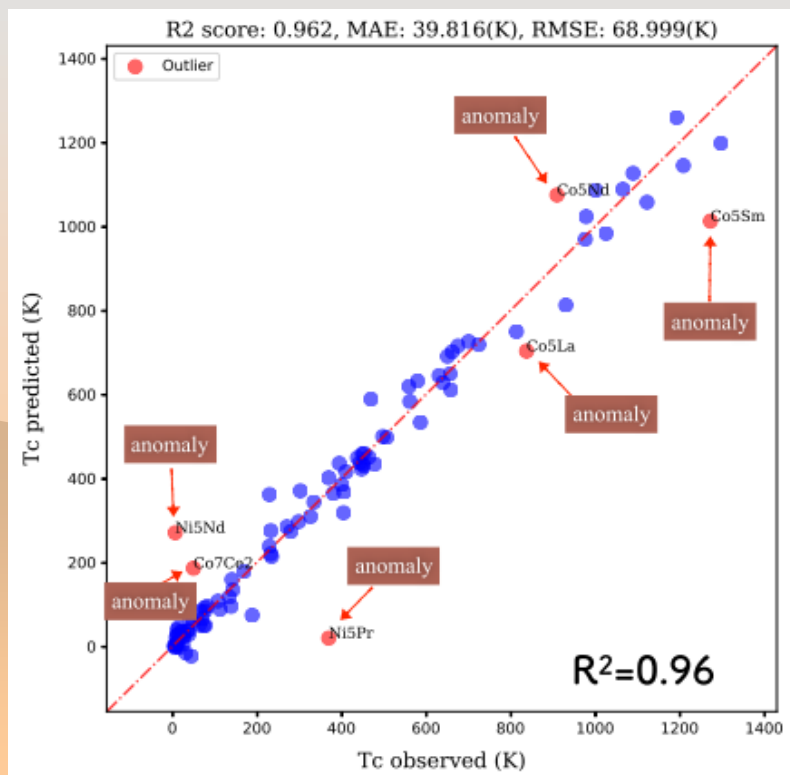
Category	Descriptors
Atomic properties of transition metals ($\bar{T}$)	$Z_T, r_T, r_T^{cv}, IP_T, \chi_T, S_{3d}, L_{3d}, J_{3d}$
Atomic properties of rare-earth metals ($\bar{R}$)	$Z_R, r_R, r_R^{cv}, IP_R, \chi_R, S_{4f}, L_{4f}, J_{4f}, g_J, J_{4f}g_J, J_{4f}(1-g_J)$
Structural information ($\bar{S}$)	$C_T, C_R, d_{T-T}, d_{T-R}, d_{R-R}, N_{T-T}, N_{T-R}, N_{R-R}, N_{R-T}$

After LASSO clustering
 R2 score: 0.970, MAE: 42.035(K)



KernelRidge回帰全探索による Subgroup relevance 解析(木野さんの方法) by 木野さん

- 108 binary alloyのCurie温度の(単一データ空間)Kernel Ridge回帰をする。
- 27原子、構造記述子を用いて全探索を行う。
- モデルの全順序リストを表示するのではなく、置き換え可能な記述子グループを定義、評価する。



TCと矛盾する記述子があると性能が落ちる。

予測モデル
を得る

data-drive
n method

TC観測値
を得る

最適化モデル
を得る

前提: 全探索が最も良い記述子セットを与える

性能最適化法

運用最適化法

最適化モデルと絡めた記述子の重要度が評価不能

解決法

linearをKernelに変える

基本記述子では回帰性能が悪い。

regression
modelを選択する

data segment
regression model分
離選択法

linear

Kernel

記述子をふくめた線形モ
デルを選択する

記述子をふくめたKernelモ
デルを選択する

global

piecewise

global modelを選択する

piecewise modelを選択する

性能最適化モ
デルを得る

性能最適化model
評価法

モデル評価値
を選択する

R^2

Dam's SCOR
E

the best

性能最適化モ
デルを選択する

運用最適化model
評価法

運用最適化モ
デルを選択する

全順序list法

記述子重要度評価法

記述子重要度評
価値を選択する

重要な記述子
sを選択する

relevance analysis

subgroup r
elevance a
nalysis

置換え可能記
述子群を同定する

記述子 subgroup
を作る

重要な
subgroupを
選択する

subgroup記述子間の
重要度を評価する

the best

運用最適化記述子の評価基準が明確でない。

両者を対比しやすいように「表現力増強記述子を得る」を作った。

TCが小さい部分の回帰に失敗していても全体の評価がよくなる。

Tc空間でのclusteringであり新規データに対する予測は難しい

subgroupの定義がない。

解決法

解決法

解決法

解決法

解決法

解決法

解決法

解決法

解決法

解決法

むすび

- 「分かる」を中心にAIを概観
- オントロジー工学は「~~言葉~~」ではなく「**概念**」を議論
- 磁石材料設計知識を対象として機能分解木の効用を確認した
 - 書くことによる知識の明確化
 - 隠れた中間ゴールの抽出
 - どんな知識が存在しているかが**分かる**
 - 知識の一覧性, 相互の関係の明確化
 - 知識の統合と体系化への一歩



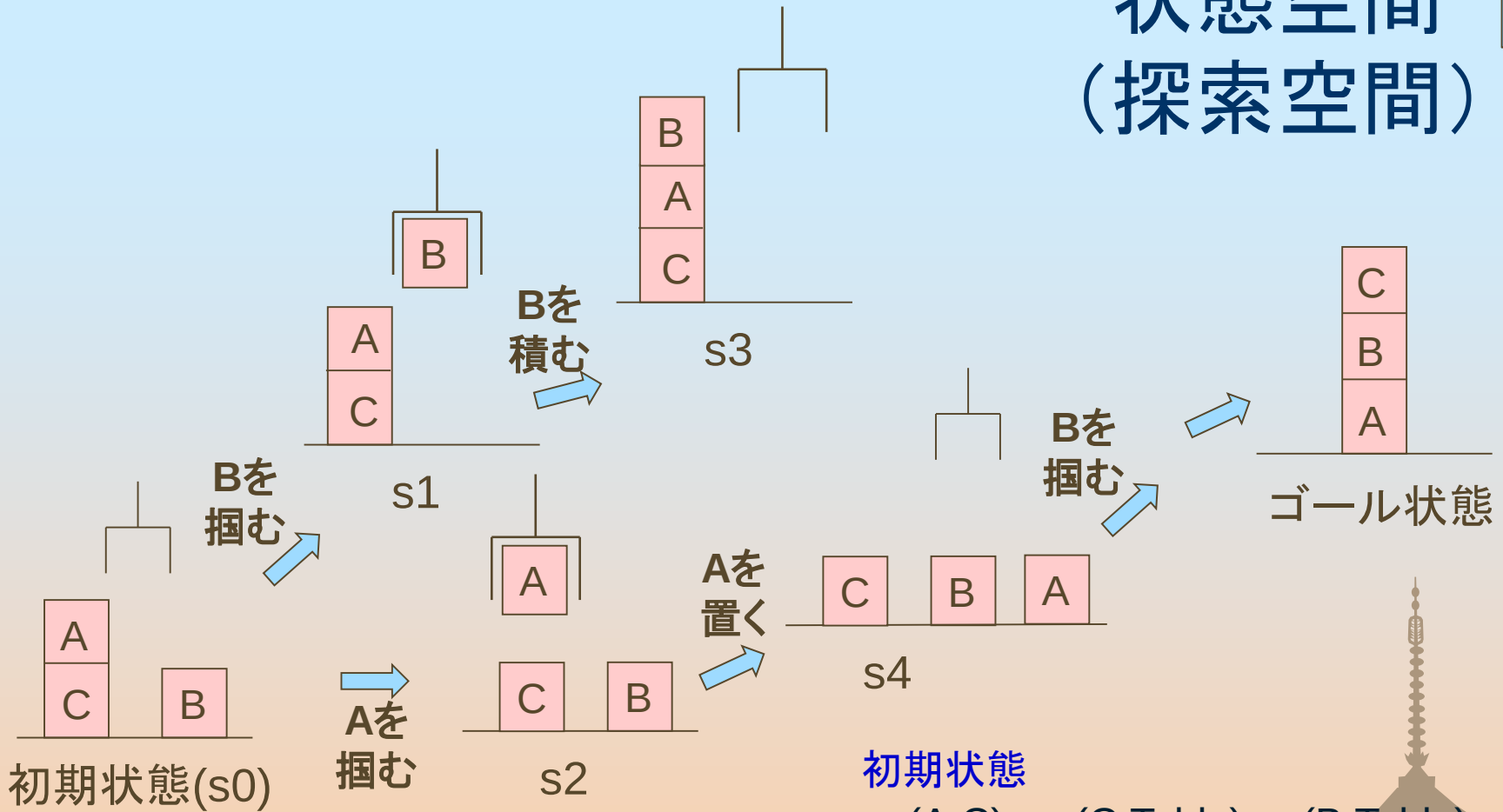
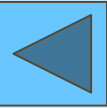


行為の分類

- 結果重視(状態変化)
 - 得る, 分ける, 殺す, 座る, 拡大する, 伝える
- 過程重視
 - 歌う, 散歩する, 付き合う, (夫婦が)会話する
- 両方
 - 歩く, 支える, なでる



状態空間 (探索空間)



掴む＝

if on(X,Y), clear(X), handEmpty
then clear(Y), holding(X)

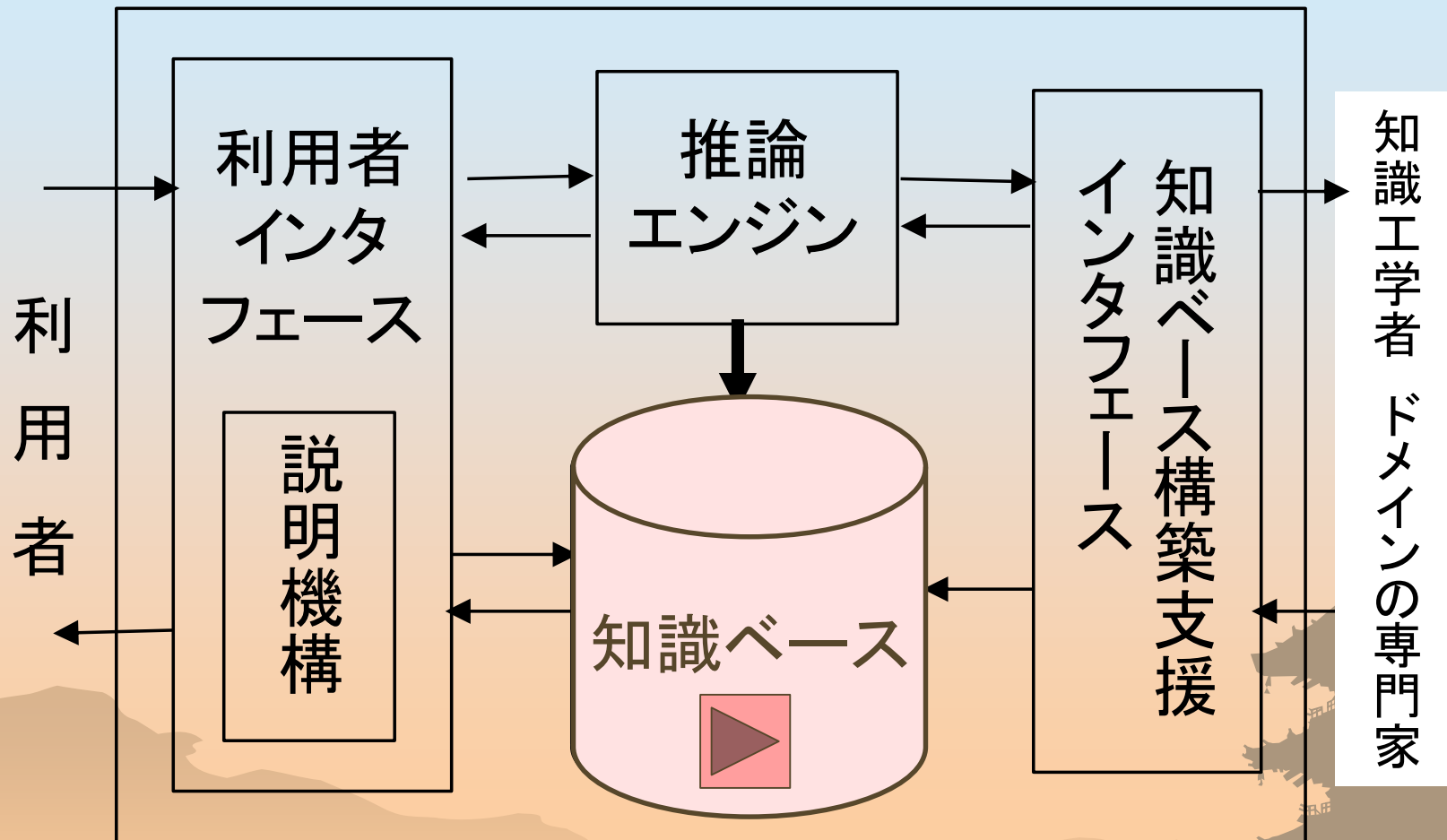
初期状態

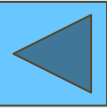
on(A,C), on(C,Table),on(B,Table)
clear(A), clear(B)
handEmpty

Aをつかんだ後

on(C,Table),on(B,Table)
clear(B),clear(C)
holding(A)

知識ベースシステムのアーキテクチャ





知識(ルール)ベースの知識の例

- If 熱が高くて, 咳が出る then 風邪を疑え
- If 振動する, 煙が出る then 部品Aの温度を測れ
- If 映画を見る then 入場料を払わなければならない
- If 暖める then 温度が上がる

専門家が持つ, 条件と結論(動作)の結合の集合体

このいい加減さをなくすために**オントロジー工学**が生まれた。存在そのものの考察から知識表現を支える理論と技術を提供する

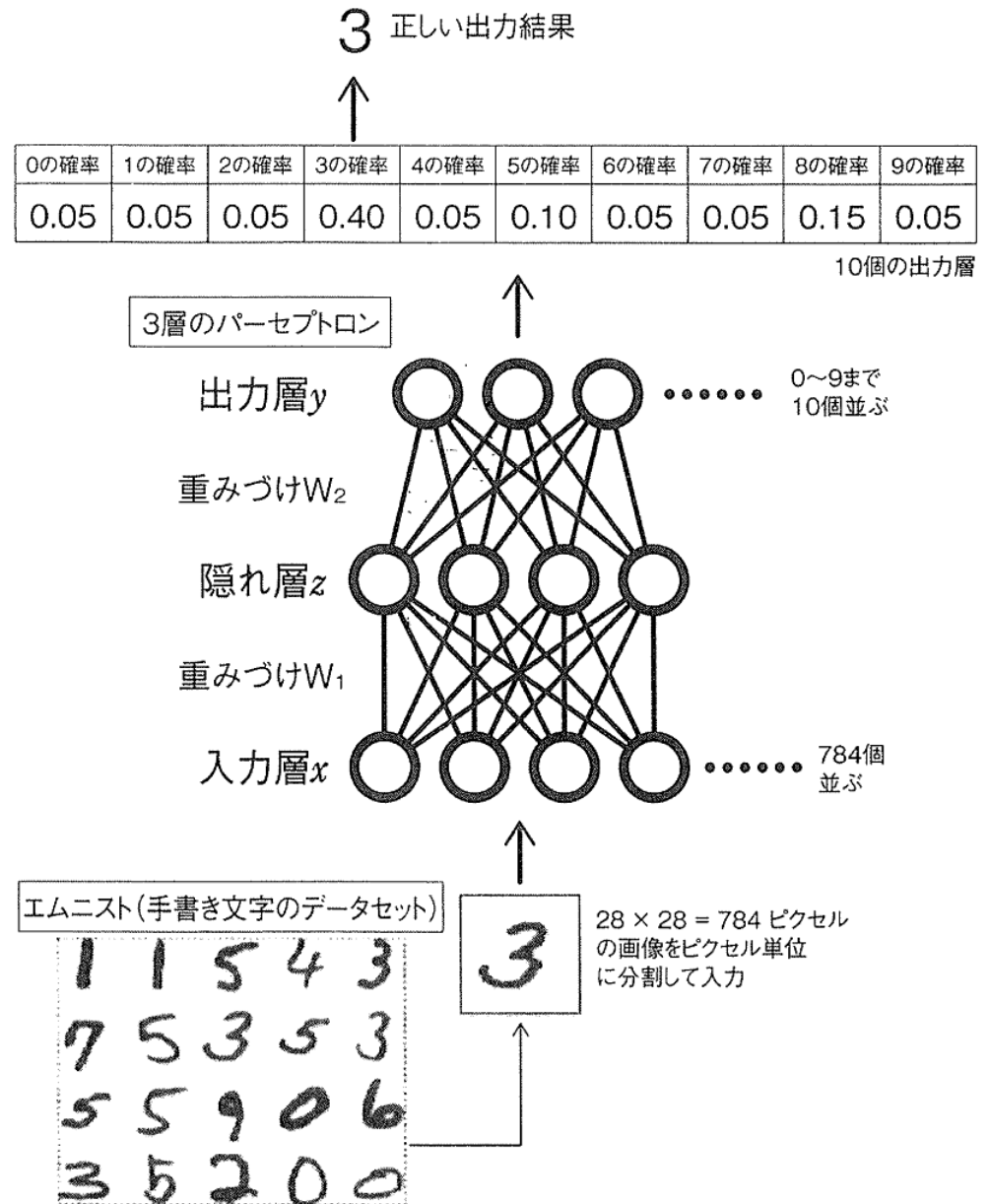


ニューラルネットワーク の構造

松尾 豊著

「人工知能は人間を超えるか」
から引用

図16 手書き文字認識



Autoencoding

自己符号化

松尾 豊著

「人工知能は人間を超えるか」
から引用

