

44年の推論研究でわかったこと

44 Years of My Reasoning Research

国立情報学研究所・情報学プリンシプル研究系

井上 克巳

Katsumi Inoue

2025年3月27日

国立情報学研究所退職記念講演

推論という研究分野

- “AI = Thinking Machine”: Can machines think? (Alan Turing)
- 1950年代から1980年代まで、推論はAI研究の主流 😊
- 1990年代以降「冬の時代」と呼ばれた時代はAI自体がタブー 😞
- 2010年代に深層学習による新たなAIブーム 😞
- 2023年より生成AIブーム 😲
- 2023年7月とあるセミナーにて「生成AI時代に求められる推論技術」
- それから2年経たないうちに 😲

今、推論()が熱い！

- **Large Reasoning Models (LRMs)** (year: 2025)

OpenAI o1, o3-mini, DeepSeek R1, Claude 3.7 Sonnet, Alibaba QwQ-32B, xAI Groc 3, etc.

- Xu, F. et al.: **Towards Large Reasoning Models: A Survey of Reinforced Reasoning with Large Language Models** (Jan 23)
- Fahey, J.: **Large Reasoning Models (LRMs): An Overview** (Jan 29)
- Li et al.: **From System 1 to System 2: A Survey of Reasoning Large Language Models** (Feb 24)
- OpenAI: **Competitive Programming with Large Reasoning Models** (Feb 18)
- Tim, D.: **The Dawn of Large Reasoning Models: AI's Quantum Leap Forward** (Feb 28)
- Huang et al.: **Safety Tax: Safety Alignment Makes Your Large Reasoning Models Less Reasonable** (Mar 1)
- Yin, H. et al.: **Towards Widening the Distillation Bottleneck for Reasoning Models** (Mar 3)
- Smith, M.S.: **It's Not Just Us: AI Models Struggle With Overthinking** (Mar 5)
- Sui, Y. et al.: **Stop Overthinking: A Survey on Efficient Reasoning for Large Language Models** (Mar 20)
- Mitchell, M.: **Artificial Intelligence Learns to Reason** (Mar 20)

自動推論(Automated Reasoning)の幕開け

- Martin Davis: Presburger's decision procedure (1954)
整数の加法演算に関する論理式 (ex. 2偶数の和は偶数) の証明
 - Newell, Shaw & Simon: "Logic Theorist" program (1955)
命題論理の置換・前向き推論・後向き推論を使って、自動的に
"Principia Mathematica" (by Alfred North Whitehead & Bertrand Russell)
第2章にある52の定理のうち38を証明、うち一つは新証明
 - Davis & Putnam: Satisfiability problem (1958)
 - 1960-1970s: 自動定理証明研究の隆盛
- Davis, M.: The Early History of Automated Deduction, in: A. Robinson & A. Voronkov:
Handbook of Automated Reasoning (2001)

AIにおける（論理的）推論 (Reasoning)

- 演繹推論 (deductive reasoning)
- 帰納推論 (inductive reasoning)
- 発想推論（アブダクション） (abductive reasoning)
- 類推 (analogical reasoning)
- 常識推論 (commonsense reasoning)
- 非単調推論 (nonmonotonic reasoning)
- デフォルト推論 (default reasoning)
- 制約推論 (constraint reasoning)
- 分散推論 (distributed reasoning)
- 確率推論 (probabilistic reasoning)
- メタ推論 (meta-reasoning)
- 事例推論 (case-based reasoning)
- 定性推論 (qualitative reasoning)
- 行動と変化に関する推論 (reasoning about action and change)
- プランニング (planning)
- 時空間推論 (spatiotemporal reasoning)
- 知識と信念に関する推論 (reasoning about knowledge and belief)
- 因果推論 (causal reasoning)
- etc.

➤ 井上克巳, 坂間千秋, 佐藤泰介: 推論とAIプログラミング, 人工知能学会誌, 34(5):703-713 (2019).

AI（機械学習/LLM）における推論 (Inference)

- 入力／質問に対し、機械学習モデルが出力／回答を返す時、AIが「推論」したともいう
- モデルへの入力：データ、条件、質問文、プロンプト、etc.
- モデルの出力：分類、予測、意思決定、行動、回答、etc.
- モデルはブラックボックスであることもあるが、出力に対する説明は重要である（XAI）
- LLMは常識推論の能力を一部備えているともいえる
- LLMでは複雑なQAタスクを複数の単純なサブタスクに置き換えて解いており (CoTなど)、これを Reasoning（多段の推論）と呼ぶことが多い



LRM で扱われる “Reasoning”

- アーキテクチャ
 - (Dense) Transformers
 - Fine-tuned for domains
 - Mixture of Experts (MoE)
 - Distilled models
 - Chain-of-Thought Self-Verification
 - Reinforcement Learning (RLHF)
 - Macro Action
 - etc.
- STEM (Science, Technology, Engineering, Mathematics) 分野
 - AIME, IMO, MATH: 数学
 - Codeforces: コード生成
 - GPQA \diamond : 博士レベル科学知識
 - 論理パズル: 命題論理で解ける問題
- 本来の（論理的な）意味での（演繹）推論に近づいた
- ただし論理パズルは不正解が多い
- 論理に従っているわけではない

これまでの研究経緯 [所属]

1981.04: 探索 (search) [京都大学]

1984.04: エキスパートシステム [松下電器産業]

1986.10: 仮説推論・非単調推論 [新世代コンピュータ技術開発機構 (ICOT)]

1993.04: + アブダクション／論理プログラミング [豊橋技術科学大学]

1997.04: + SAT／帰納論理プログラミング [神戸大学]

2004.03: + マルチエージェントシステム [国立情報学研究所 (NII)]

2007頃-: + システム生物学

2012頃-: + レジリエンス／ダイナミクス学習

2017頃-: + 知識表現・推論と機械学習の融合

以上すべて、推論(reasoning)に関係

ゲーム木探索

(茨木俊秀先生による指導)

- 大学4年で初めて本格的に読んだ論文で、ゲーム木 (Minimax木) を解く α - β アルゴリズムを理解
- ゲーム木を解くヒューリスティック探索アルゴリズムの研究：卒論・修論研究において、既存探索手法の性能を上回るアルゴリズムの開発を目指し実現した
- ゲーム木のみならず、一般の探索における枝刈り手法を習得した
- 何より研究に対する姿勢を学んだ



Artificial Intelligence
Volume 6, Issue 4, Winter 1975, Pages 293-326



An analysis of alpha-beta pruning ☆

Donald E. Knuth, Ronald W. Moore

Show more ▾

+ Add to Mendeley  Share  Cite

[https://doi.org/10.1016/0004-3702\(75\)90019-3](https://doi.org/10.1016/0004-3702(75)90019-3)

[Get rights and content ↗](#)

Abstract

The alpha-beta technique for searching game trees is analyzed, in an attempt to provide some insight into its behavior. The first portion of this paper is an expository presentation of the method together with a proof of its correctness and a historical discussion. The alpha-beta procedure is shown to be optimal in a certain sense, and bounds are obtained for its running time with various kinds of random data.

探索の効率化は知性か？

- “Intelligence = Search”
- 導出原理に基づく定理証明戦略はゲーム木探索に似た発見法 [1]
(AND/OR木探索 \approx Minimax木探索)
- 問題解決は宝探しであり、宝をいかに賢く探し当てるか [2]
- 後に、仮説推論・結論発見・制約充足・帰納推論など多くの分野で探索が重要であると認識し、研究内容に活かすことができた

[1] James R. Slagle: *Artificial Intelligence: The Heuristic Programming Approach* (1971)

[2] Herbert A. Simon: Search and Reasoning in Problem Solving, *Artificial Intelligence*, 27 (1983)

エキスパートシステム

- 1980年代に流行し、AIブームを牽引
- **プロダクションルール**: 専門家がもつ手続き的知識を“IF-THEN”形式のルールで表現。ルールのIF部（条件部）が現在の状態で満たされるとそのルールは発火し(**fire**)、ルールのTHEN部（実行部）が導かれる
- **黒板モデル**(blackboard model)に基づく推論エンジン（シェル）をLISPで設計し、社内事業部と共同で金型工程设计エキスパートシステムを開発
- この過程で**設計のための推論原理**に興味を抱き基礎理論の重要性を認識

THURSDAY, NOVEMBER 17, 1-3 PM

EDWARD A. FEIGENBAUM

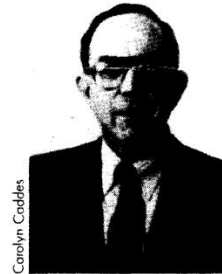
AND

H. PENNY NII

Will sign copies of

The Rise of the Expert Company

How Visionary Companies are Using Artificial Intelligence to Achieve Higher Productivity and Profits



Carolyn Caddles

Edward Feigenbaum is Professor of Computer Science at Stanford University and is Scientific Director of the Heuristic Programming Project, also at Stanford. He is a co-founder of two start-up firms in applied artificial intelligence, Intellicorp and Teknowledge; and serves as Chairman of the board of directors of Intellicorp.



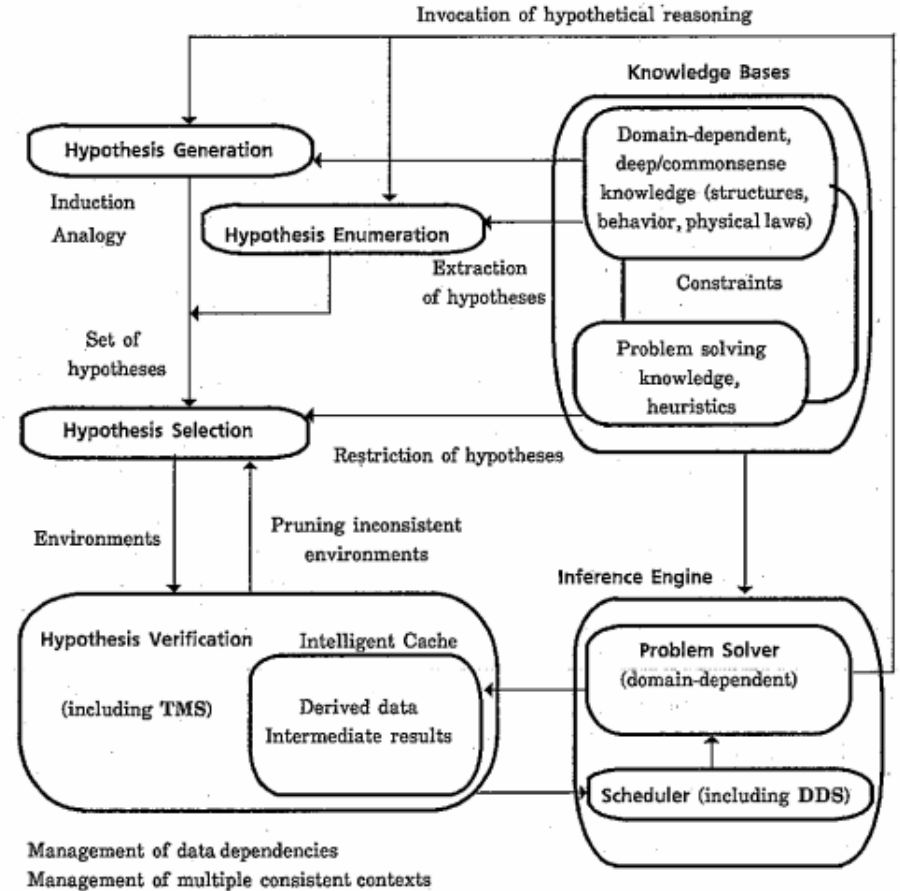
Carolyn Caddles

H. Penny Nii, a senior research scientist at Stanford University, was one of the first “knowledge engineers” working on the delicate task of transforming human expertise into the novel computer data structures eventually known as expert systems.

仮説推論

(Hypothetical Reasoning)

- 問題解決の途中で仮説を立てて推論し、解けない時は別の仮説を試す (コンテキストの変更)
- 矛盾する仮説の組合せ(**nogood**)は後にCSP/SAT効率化の基礎をなす [1]
- 並列論理型言語KL1で仮説に基づく真偽維持システム(ATMS)および仮説推論システムAPRICOTを実現し [2]、富士通社とともにVLSI設計に応用
- これらの技術は当時、高次推論 (advanced reasoning)と呼ばれたが、論理的な仕組みが不明であった



[1] 井上克巳: SATとAI. 情報処理, 57(8) (2016)

[2] Katsumi Inoue: Problem Solving with Hypothetical Reasoning. FGCS 1988

仮説推論は論理で表現できる！

- ATMSはprime implicate（主節）計算と関係 (Reiter & de Kleer, AAAI-87)
- 主節を計算することで命題論理上のアブダクションが実現できる
- 命題論理説集合からの結論発見が使えることがわかった

Example 5.3 Let an ATMS be $\langle \{a, b, c, x, \neg y\}, \{x, \neg y\}, \Sigma \rangle$ where

$$\Sigma = \left\{ \begin{array}{ll} \neg a \vee \neg b \vee c & (1), \\ \neg x \vee \neg b \vee a & (2), \\ y \vee b \vee c & (3) \end{array} \right\}.$$

The following deduction finds c 's label $\{x \wedge \neg y\}$:

$\langle \square, \underline{\neg c} \rangle,$	top clause
$\langle \square, \underline{\neg a} \vee \neg b \vee \boxed{\neg c} \rangle,$	resolution with (1)
$\langle \square, \underline{\neg x} \vee \cancel{\neg b} \vee \boxed{\neg a} \vee \neg b \vee \boxed{\neg c} \rangle,$	resolution with (2) and factoring
$\langle \neg x, \boxed{\cancel{\neg d}} \vee \neg b \vee \boxed{\neg c} \rangle,$	skip and truncation
$\langle \neg x, y \vee \cancel{c} \vee \boxed{\neg b} \vee \boxed{\neg c} \rangle,$	resolution with (3) and reduction
$\langle \neg x \vee y, \boxed{\cancel{\neg d}} \vee \boxed{\cancel{\neg d}} \rangle.$	skip and truncation

結論発見 (consequence finding)

- 演繹推論であっても、知られていない定理を生成すれば発見と言える
- 面白い定理とは何か？それらをどのように生成できるか？
- それまで未解決であったモデル消去法(効率的な線形導出)の枠組における結論発見手続きを提案し、その完全性を証明
- 生成領域(production field)という概念を導入し、定理の「面白さ」を定義
- 手続きを命題論理から一階述語論理に拡張
- 初めてIJCAI (1991)に論文が採択され、その拡張版がAI Journalに掲載された



Linear resolution for consequence finding

Katsumi Inoue  

Show more 

 Add to Mendeley  Share  Cite

[https://doi.org/10.1016/0004-3702\(92\)90030-2](https://doi.org/10.1016/0004-3702(92)90030-2)

[Get rights and content](#) 

Abstract

In this paper, we re-evaluate the consequence finding problem within first-order logic. Firstly, consequence finding is generalized to the problem in which only interesting clauses having a certain property (called characteristic clauses) should be found. The use of characteristic clauses enables characterization of various reasoning problems of interest to AI, including abduction, nonmonotonic reasoning, prime implicates and truth maintenance systems. Secondly, an extension of the Model Elimination theorem proving procedure (SOL-resolution)

節理論における結論発見法

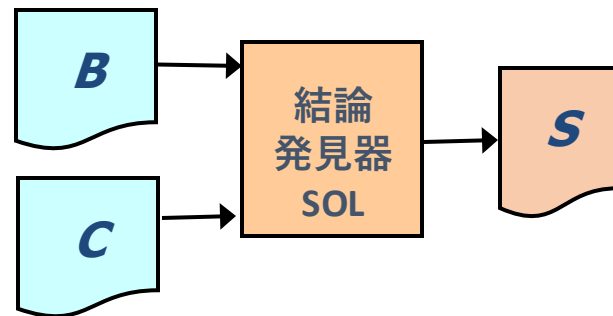
入力 B : 背景知識 (一階述語節理論)

C : 追加される知識 (一階述語節理論)

P : 言語バイアス (生成領域, *production field*)

出力 S : 以下を満たす極小節:

- $B \cup C \models S$
- $B \not\models S$
- S は P に属する



● **SOL導出** (Inoue, AIJ 1992)

● **SOLAR** (Nabeshima, Iwanuma & Inoue, TABLEAUX 2003; AIC 2010)

● AIの各種問題への応用: 定理証明($P = \{\}$)、条件付き質問応答 (JELIA 2002)、**アブダクション** (AIJ 1992)、**インダクション** (MLJ 2004)、デフォルト推論 (JIIS 2004)、マルチエージェントシステム (AMAI 2004)、**システム生物学**、etc.

群論に関する定理発見

$$C1: p(e, x, x)$$

$$C2: p(i(x), x, e)$$

$$C3: p(x, y, u) \wedge p(y, z, v) \wedge p(u, z, w) \rightarrow p(x, v, w)$$

$$C4: p(x, y, u) \wedge p(y, z, v) \wedge p(x, v, w) \rightarrow p(u, z, w)$$

$$C5: p(x, i(x), e)$$

$$C6: p(x, e, x)$$

- $(e \circ x) = x$ [左単位元]
- $(i(x) \circ x) = e$ [左逆元]
- $(x \circ y) \circ z = w \rightarrow x \circ (y \circ z) = w$ [結合律]
- $x \circ (y \circ z) = w \rightarrow (x \circ y) \circ z = w$ [結合律]

-
- $(x \circ i(x)) = e$ [左逆元は右逆元でもある]
 - $(x \circ e) = x$ [左単位元は右単位元でもある]

$B = \{C1, C2, C3\}, C = \{C4\}$ とすると, $B \cup C \models_P C5 \wedge C6$
生成領域: $P = \langle \{p\}^+, \text{length} \leq 1 \ \& \ \text{term_depth} \leq 1 \rangle$

アブダクション・インダクションの論理

入力:

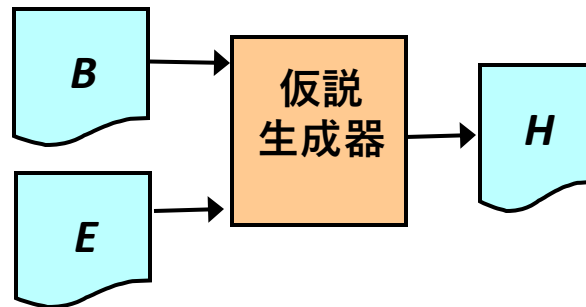
B : 背景知識

E : 観測・事例

出力:

H : 次の条件を満たす**仮説**

- $B \wedge H \models E$
- $B \wedge H$ は無矛盾



逆帰結法 (Inverse Entailment; IE)

仮説 H の否定は演繹的に計算できる：

$$B \wedge \neg E \models \neg H$$

つまり、結論発見手続きを用いることができる

▶ 井上克巳: アブダクションとインダクション, 人工知能学会誌, 25(3):389-399 (2010)

CF帰納法 (Inoue, 2001) [1]

$B \wedge \neg E \models \neg H$

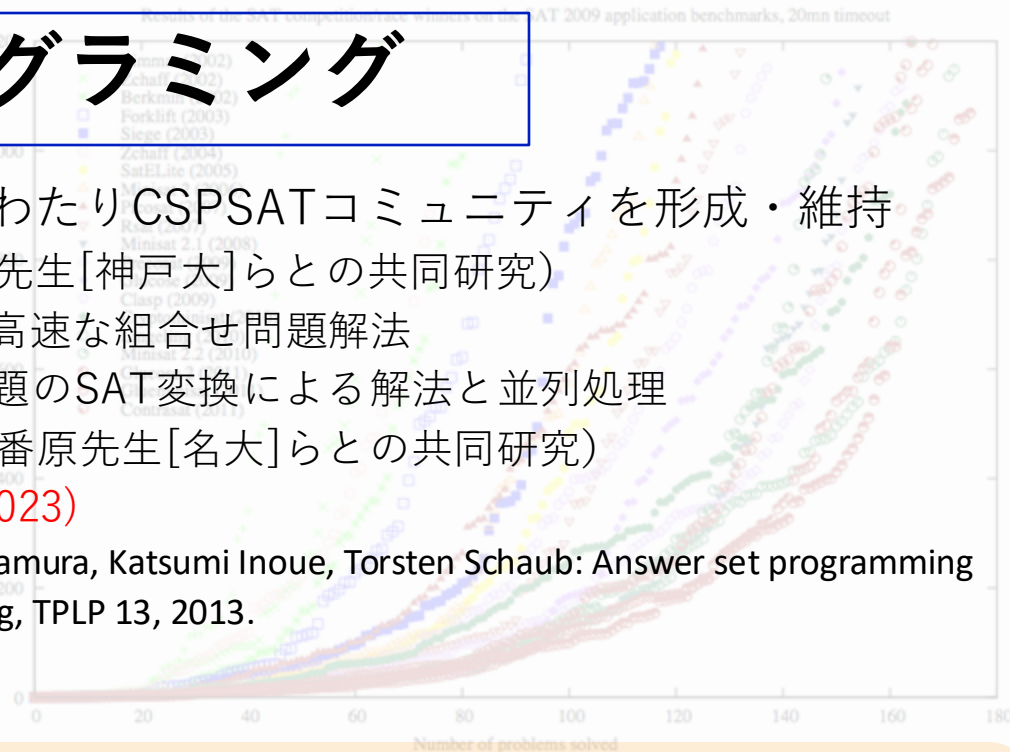
- 逆帰結法において、ルールを含む仮説 H を得る手続き
- 一階述語節理論に対して健全かつ完全
- $B \wedge \neg E$ の**特徴節**を結論発見手続きで計算
- 汎化手続きを利用した一般化
- 過去に提案されたすべての逆帰結法を包含した一般化：
 - B, E, H : いずれもフルの一階述語節理論 (非ホーン節を含む)
- 帰結(entailment)を包摂(subsumption)に置き換えることも可能 [2]

[1] Katsumi Inoue: Induction as Consequence Finding. *Machine Learning*, 55(2): 109-135 (2004)

[2] Yoshitaka Yamamoto, Katsumi Inoue, Koji Iwanuma: Inverse subsumption for complete explanatory induction. *Machine Learning*, 86(1): 115-139 (2012)

制約充足・制約プログラミング

- NII共同研究を通じて国内で20年にわたりCSPSATコミュニティを形成・維持
- SATに基づく制約モデリング（田村先生[神戸大]らとの共同研究）
 - ▶ 命題論理の充足判定(SAT)を用いた高速な組合せ問題解法
 - ▶ 制約充足問題(CSP)・制約最適化問題のSAT変換による解法と並列処理
- 解集合プログラミング(ASP)研究（番原先生[名大]らとの共同研究）
 - 10-Year Test-of-Time受賞 (ICLP 2023)
 - ▶ Mutsunori Banbara, Takehide Soh, Naoyuki Tamura, Katsumi Inoue, Torsten Schaub: Answer set programming as a modeling language for course timetabling, TPLP 13, 2013.
- 検証・設計・診断等への応用



SAT Competition 受賞ソルバー（鍋島先生[山梨大]らとの共同研究）

- glueminisat (SAT 2017): 擬似LBDによるリスタート (App UNSAT 2011 🏆🏆🏆, 2013 🏆)
- ManyGlucose (SAT 2020): 決定性を有する並列SATソルバー (2020 Parallel 🏆)
- DPS: ManyGlucoseを一般化し Kissat と接続 (2022 Main Parallel 🏆)



湧一博記念コロキウム
『論理と推論技術：四半世紀の展開』

非単調性と帰納論理を取り入れたことで
論理プログラミングはどう変わったか？

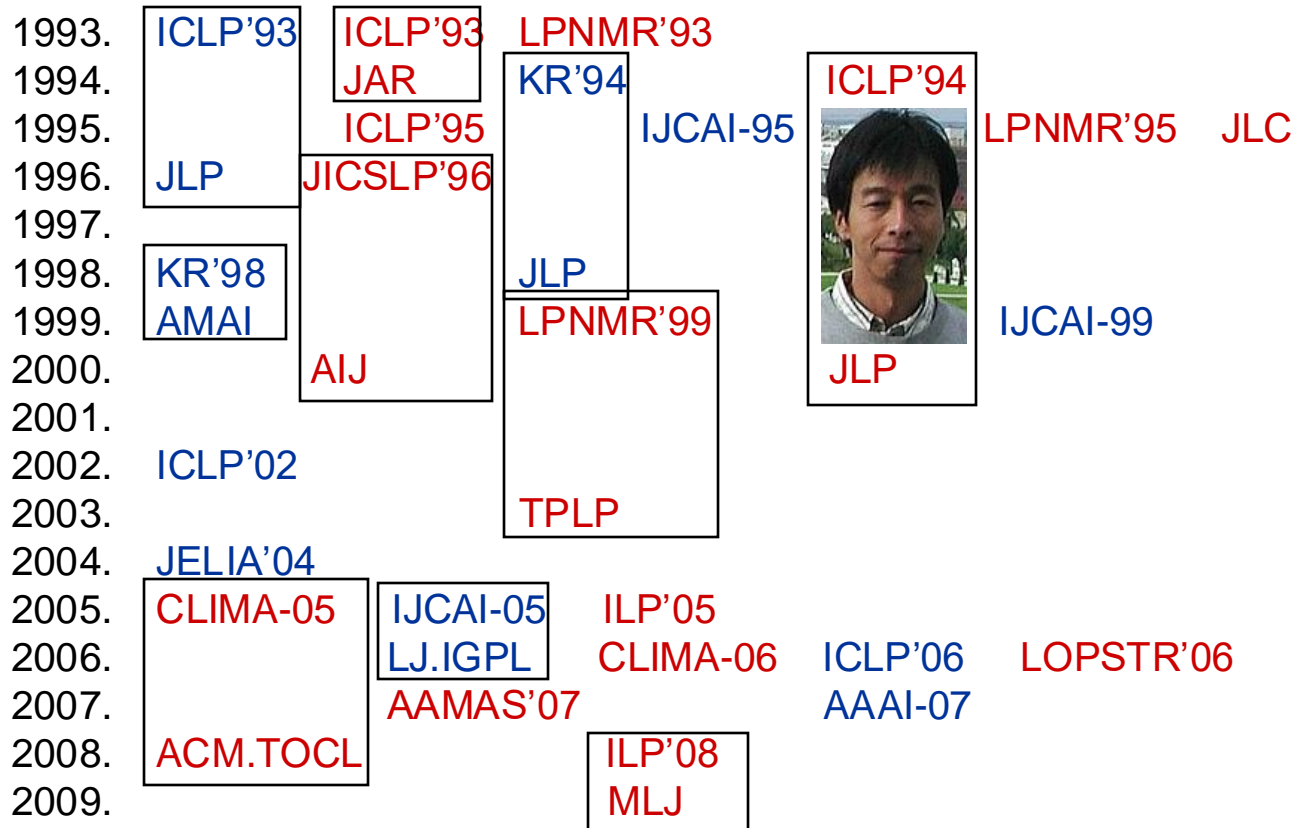
国立情報学研究所
情報学プリンシプル研究系

井上 克巳

2007年10月20日

論理プログラミング基礎理論

(Inoue & Sakama / Sakama & Inoue)



- negation
- disjunction
- circumscription
- paraconsistency
- **abduction**
- preference
- update
- negotiation
- coordination
- equivalence
- generality
- induction

アブダクション [再]

背景知識 B および観測 E が与えられたとき、
アブダクションは以下の関係を満たす
 G の説明 H を推論する：

1. $B \cup H \models E$
2. $B \cup H$ は無矛盾

➤ 井上克巳:アブダクションの原理. 人工知能学会誌 7(1):48-59 (1992)

拡張アブダクション

背景知識 B および観測 E が与えられたとき、
拡張アブダクションは以下の関係を満たす

- G の説明 (P, N) :

$$(B \setminus N) \cup P \models E$$

- または G の反説明 (P, N) :

$$(B \setminus N) \cup P \not\models E$$

を推論する。ただし $(B \setminus N) \cup P$ は無矛盾

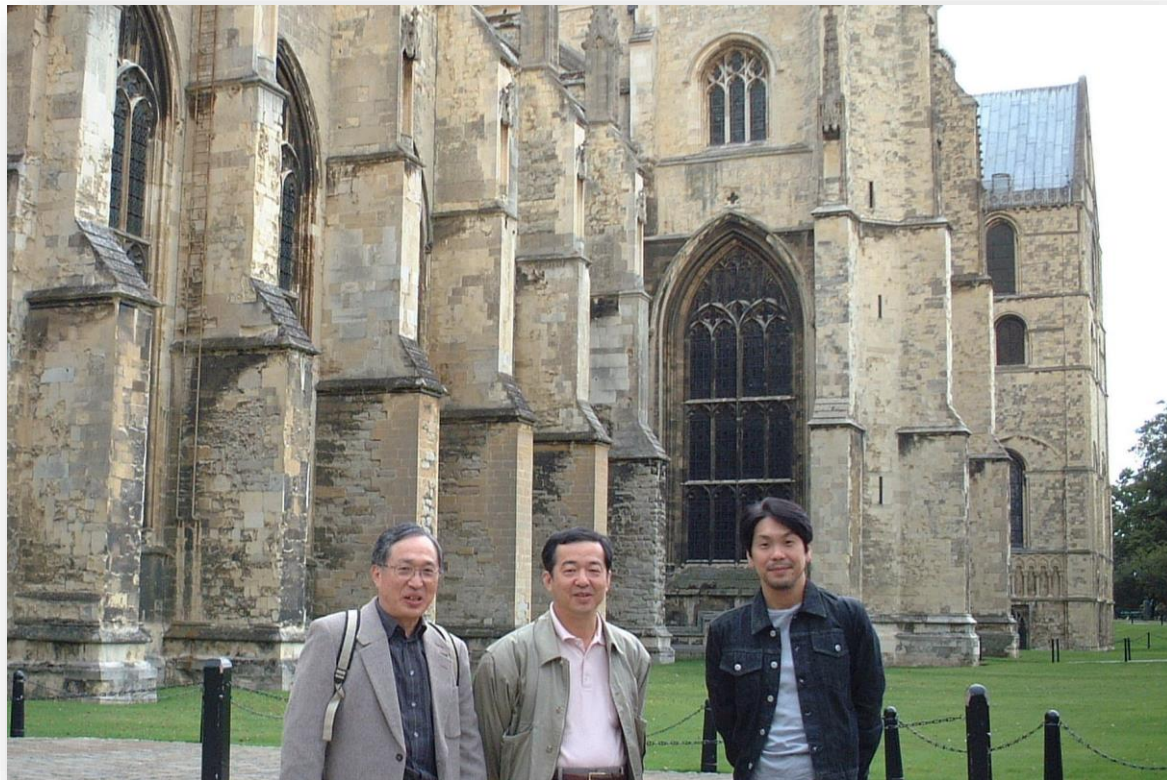
- Katsumi Inoue, Chiaki Sakama: Abductive Framework for Nonmonotonic Theory Change. IJCAI 1995
- Katsumi Inoue, Chiaki Sakama: Specifying Transactions for Extended Abduction. KR 1998

拡張アブダクションの応用

- 非単調理論におけるアブダクション (Inoue & Sakama, IJCAI-95)
- データベースにおけるビュー更新 (Sakama & Inoue, LPNMR 1999)
- 理論の修正・更新 (Inoue & Sakama, ICLP 2002)
- 矛盾除去 (Sakama & Inoue, TPLP 2003)
- システム修正・モデル検査 (*)
- ILP (inductive logic programming)
- 非単調論理プログラムにおける拡張アブダクションから通常のアブダクションへのプログラム変換 (Inoue, CL-2000)

(*) Bashar Nuseibeh & Alessandra Russo: Using Abduction to Evolve Inconsistent Requirements Specification. *Australas. J. Inf. Syst.* 6(2) (1999)

古川康一先生と佐藤泰介先生



Canterbury Cathedral, St., UK (Machine Intelligence Workshop 19), 2002

メタレベル・アブダクション

- メタプログラム上でアブダクションを行う
- 規則アブダクション (law abduction) の一手法:
 - a. 経験則を隠れ規則により説明する
 - b. 複数のミッシング・リンクや因果関係を発見する
 - c. 複数の観測事実／目標を同時に説明／実現する
 - d. 新概念を新述語として含むことができる
- 古川先生による楽器演奏技術のコツを説明する研究がきっかけ
- 負の影響を入れ、生体ネットワークデータからの科学的発見に適用
- メタ解釈学習 (MIL) (Muggleton *et al.*, 2013-) の先駆的研究



- Katsumi Inoue, Koichi Furukawa, Ikuo Kobayashi, Hidetomo Nabeshima: Discovering Rules by Meta-level Abduction, *Post-Proc. ILP 2009*, LNAI 5989, pp.49-64 (2010)
- Tomonobu Ozaki, Randy Goebel, Katsumi Inoue: From Fifth Generation Computing to Skill Science—A Biographical Essay of Koichi Furukawa, *New Generation Computing*, 37(2):141-158 (2019)

メタレベルアブダクションによる推論補完

B: $linked(increase_sound, bow_close_to_the_bridge)$.

$linked(bow_close_to_the_bridge, stable_bow_movement) \vee$

$linked(bow_close_to_the_bridge, smooth_bow_direction_change)$.

$linked(smooth_bow_direction_change, flexible_wrist)$.

$\leftarrow linked(inc_sound, keep_arm_close)$.

$\leftarrow linked(stable_bow_movement, keep_arm_close)$.

$\leftarrow linked(smooth_bow_direction_change, keep_arm_close)$.

A: $caused(X, Y) \leftarrow linked(X, Y)$.

$caused(X, Y) \leftarrow linked(X, Z) \wedge caused(Z, Y)$.

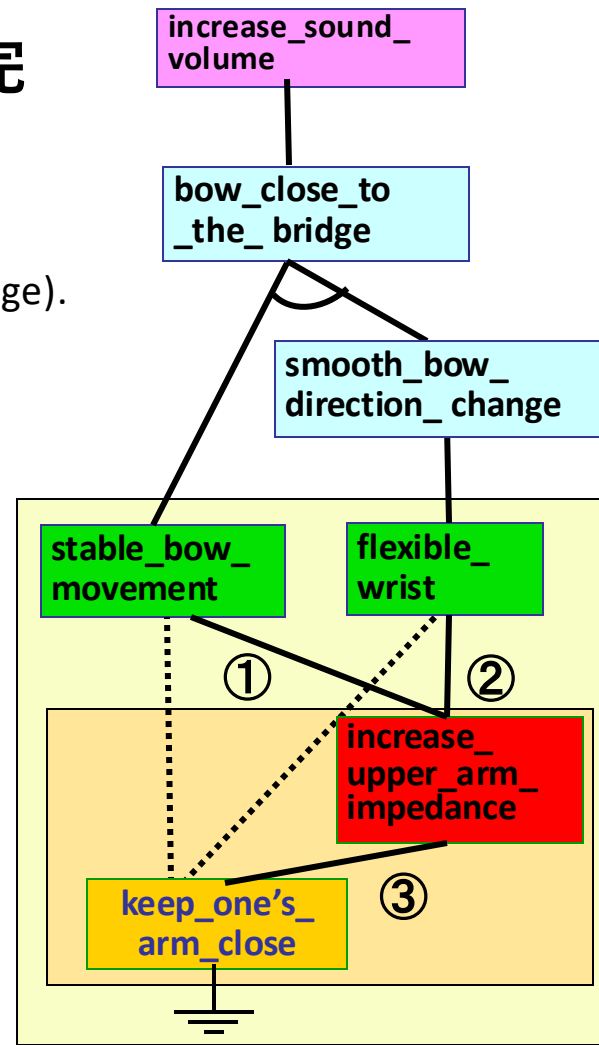
G: $\leftarrow caused(increase_sound, keep_arm_close)$.

H: SOLAR が計算した52個の仮説の一つ (Xは新述語に相当) :

$\exists X (linked(stable_bow_movement, X) \quad \textcircled{1}$

$\wedge linked(flexible_wrist, X) \quad \textcircled{2}$

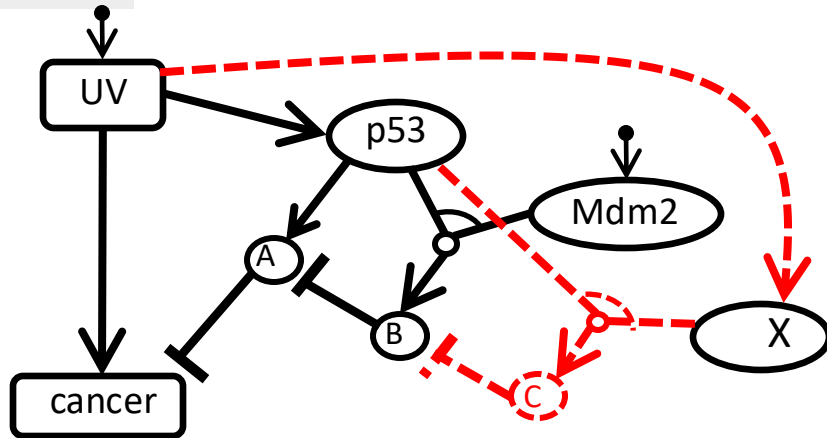
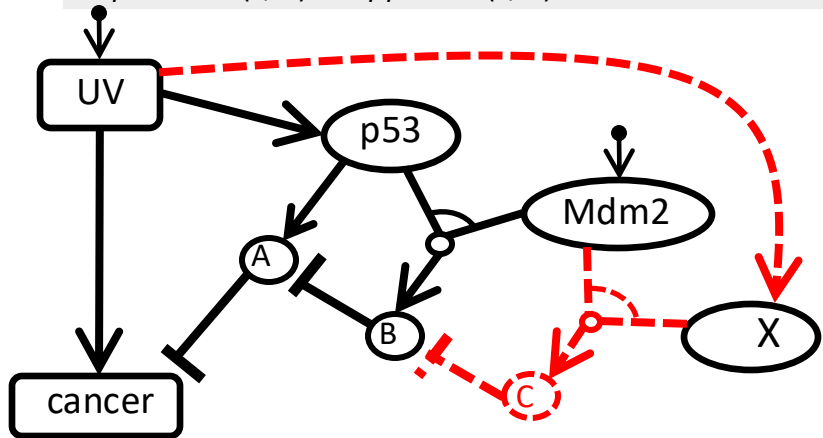
$\wedge linked(X, keep_arm_close)) . \quad \textcircled{3}$



負の効果をもつリンク補完とデフォルト推論

$promoted(X, Y) \leftarrow triggered(X, Y) \wedge no_inhibitor(X).$
 $promoted(X, Y) \leftarrow triggered(X, Z) \wedge no_inhibitor(X) \wedge promoted(Z, Y).$
 $promoted(X, Y) \leftarrow inhibited(X, Z) \wedge suppressed(Z, Y).$
 $suppressed(X, Y) \leftarrow inhibited(X, Y).$
 $suppressed(X, Y) \leftarrow inhibited(X, Z) \wedge promoted(Z, Y).$
 $suppressed(X, Y) \leftarrow triggered(X, Z) \wedge no_inhibitor(X) \wedge suppressed(Z, Y).$
 $\leftarrow promoted(X, Y) \wedge suppressed(X, Y).$

- 正負の効果をもつ因果ネットワーク
- メタレベルアブダクションにおけるデフォルト推論
- 促進因子からの影響は、抑制因子からの影響がないと仮定しても矛盾しない場合にのみ働く



$triggered(x, uv) \wedge \exists C (jointly_triggered(C, mdm2, x) \wedge inhibited(b, C))$
 $triggered(x, uv) \wedge \exists C (jointly_triggered(C, p53, x) \wedge inhibited(b, C))$

JSAI2013

2013年度人工知能学会全国大会

説明と予測： 科学的発見からレジリエンスへ

国立情報学研究所

情報学プリンシプル研究系

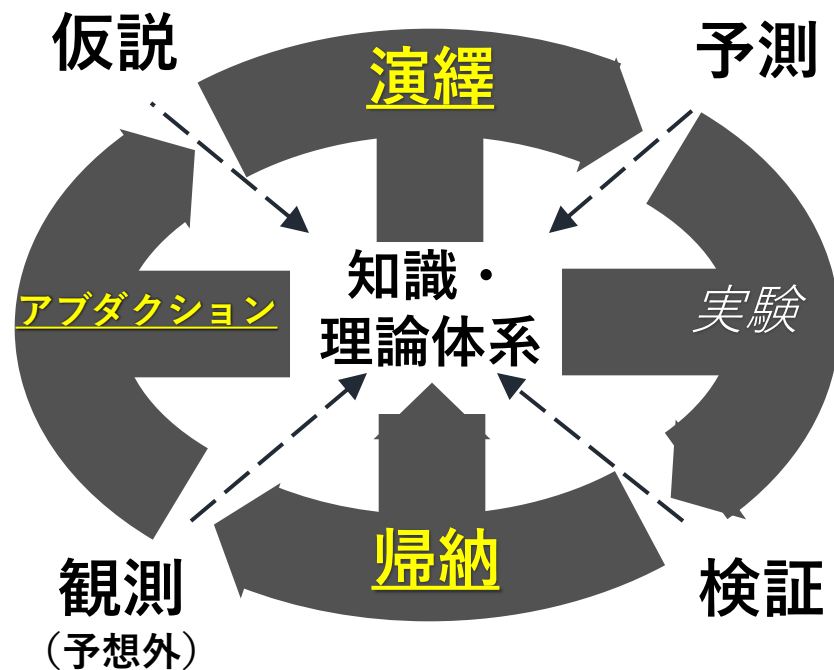
井上 克巳



国立情報学研究所
National Institute of Informatics

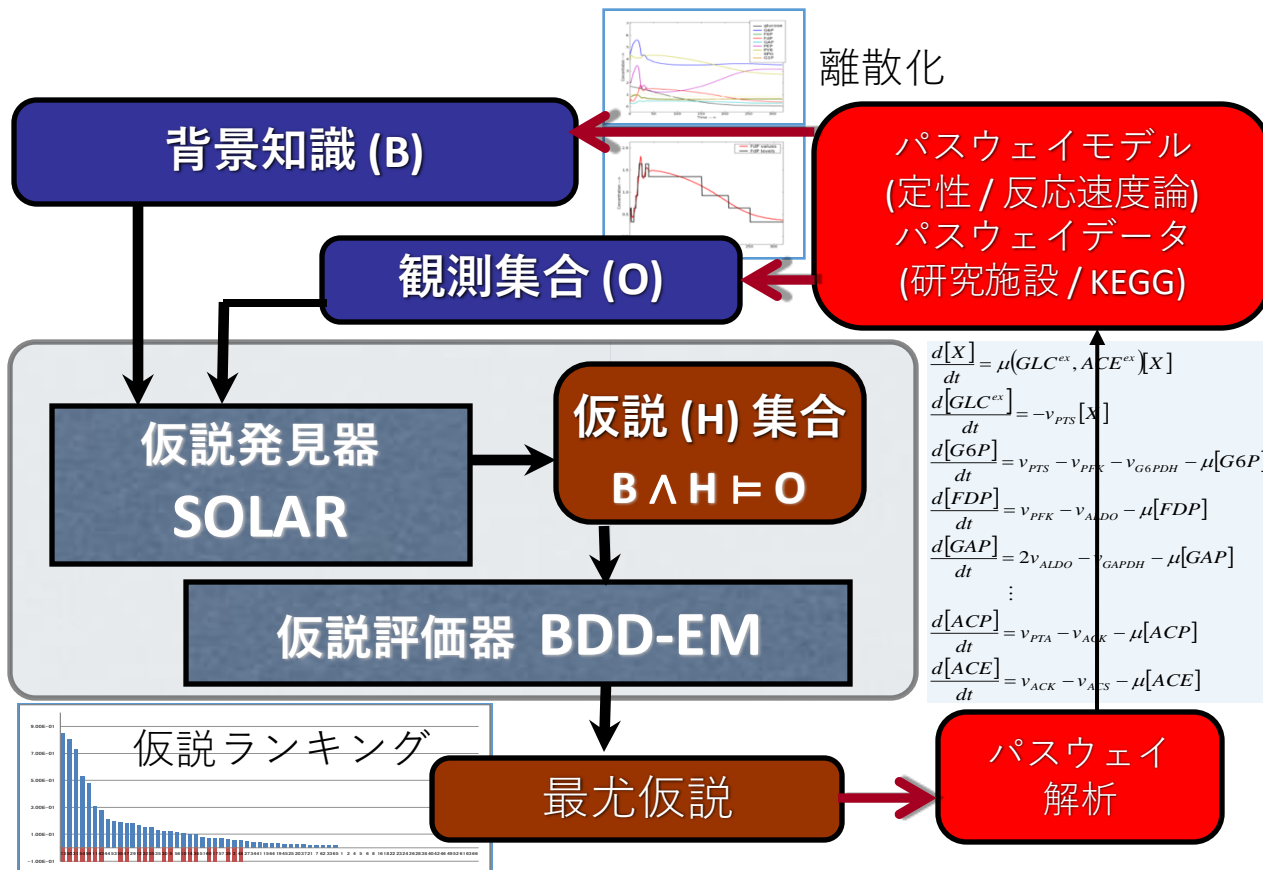
科学知識の生成・発展サイクル

- アブダクションで得られた仮説を検証し、次の実験計画を決定し、新たな観測を得て仮説を更新して行く
- アブダクション→演繹→帰納のサイクルが繰り返され、科学理論が洗練されていく
- 背景理論と理論修正が重要
- ループ構造のない帰納推論だけだと真の発見には結びつかない
- Robot Scientist (King et al., 2004-2009) は全自動でこれらの実践を目指し、アブダクションにSOL導出を利用



➤ 井上 克巳: 人工知能による科学的発見. 電子情報通信学会誌 98(1):35-39 (2015)

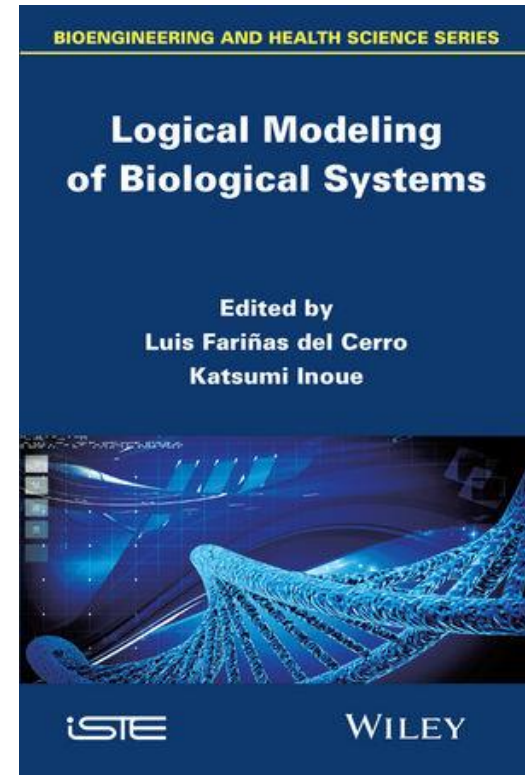
システム生物学における仮説発見・評価



- Katsumi Inoue, Taisuke Sato, Masakazu Ishihata, Yoshitaka Kameya, Hidetomo Nabeshima: Evaluating Abductive Hypotheses using an EM Algorithm on BDDs. IJCAI 2009

生物・生命の論理モデリング

- 要素間の関係を制約や関係述語として宣言的に記述
- 因果関係の第一近似として系の定性的性質を推論
- 論理的説明により専門家の理解が容易
- 背景理論（既存生物学知識・DB）を伴った学習
- 不完全理論における理論補完
- 論理的に可能な複数仮説の枚挙
- 能動的学習による実験計画作成
- シミュレーションに代わるモデル検査法
- 確率と論理の結合による推論・学習ツール
- 遺伝子／代謝物質の相互関係ネットワークを論理表現
- ネットワーク補完：アブダクションによるミッシングリンク発見



➤ Luis Fariñas del Cerro, Katsumi Inoue (eds.): *Logical Modeling of Biological Systems*, Wiley (2014)

喫緊の課題

- 2011年3月：東日本大震災
- 2020年3月：COVID-19パンデミック
- 非常に困難な状況に直面した時、研究者として、あるいは一人の人間として、自分は何が出来るだろうか？
 - 丸山宏, 井上克巳, 椿広計, 明石裕, 岡田仁志, 南和宏: システムズ・レジリエンス, 第11回情報科学技術フォーラム (FIT 2012)
 - Moshe Vardi: “Efficiency vs. resilience: what COVID-19 teaches computing”, *Communications of the ACM*, 63(5):9 (2020)

システムズ・レジリエンス (2012-2016)

1. 震災等の想定外事象が起きた際に、擾乱に対してロバストであり、損失から早期に機能回復できる系の性質はどのようなものか？

- 耐性 (ロバスト性)
- 冗長性・分散系
- 多様性・多機能性
- 適応性



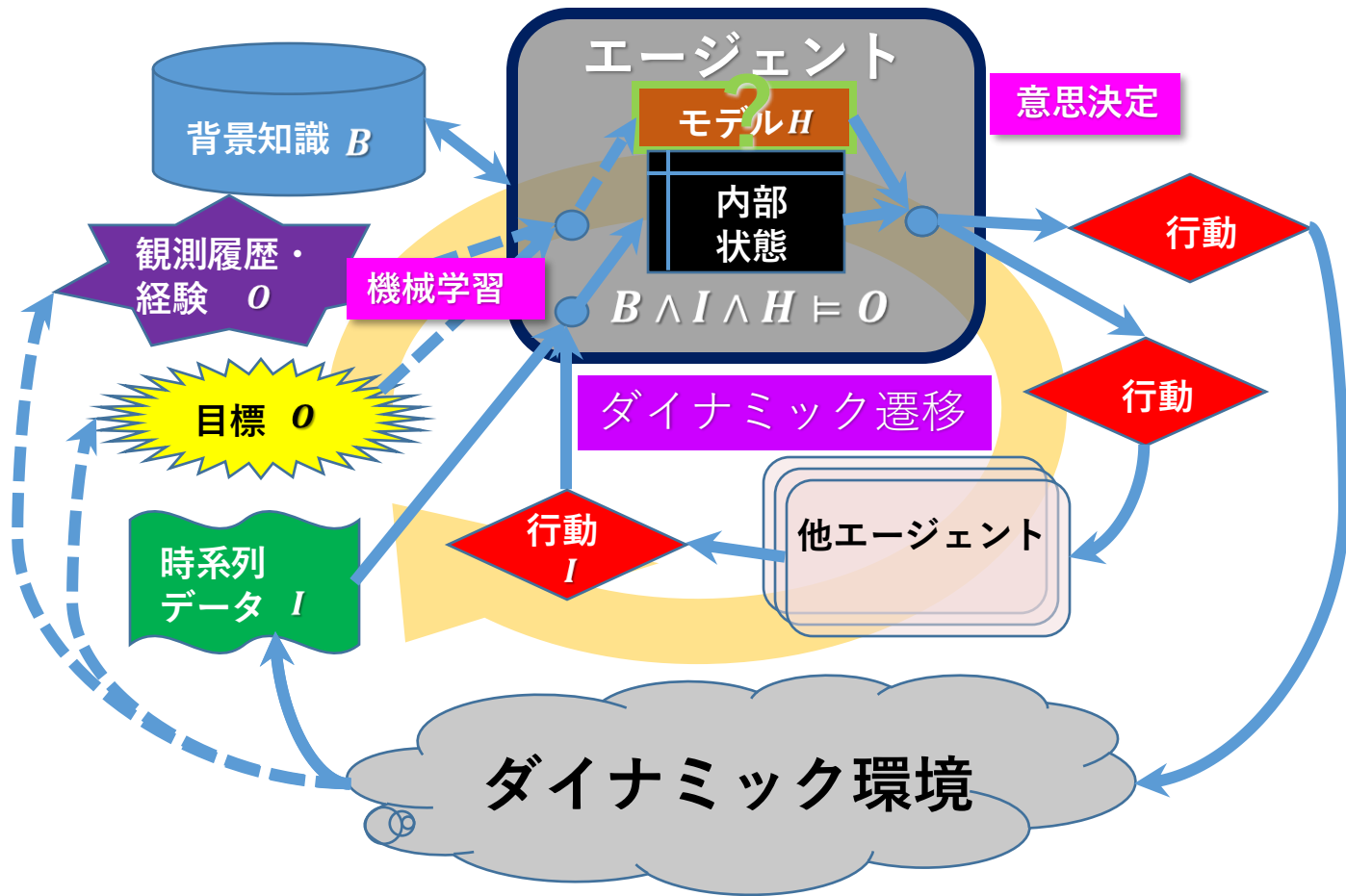
2. レジリエントな系の設計～人工知能による貢献

- 記号表現—ダイナミック制約ネットワーク
- 予測・説明—予測 (不) 可能性の論理
- 多目的制約最適化の理論・実装・応用
- ロバストな系の設計・解析・計算量理論→ロバストAI
- アトラクター—遷移

➤ Nicolas Schwind, Morgan Magnin, Katsumi Inoue, Tenda Okimoto, Taisuke Sato, Kazuhiro Minami, Hiroshi Maruyama: Formalization of resilience for constraint-based dynamic systems, *J. Reliable Intelligent Environments*, 2(1):17-35 (2016)

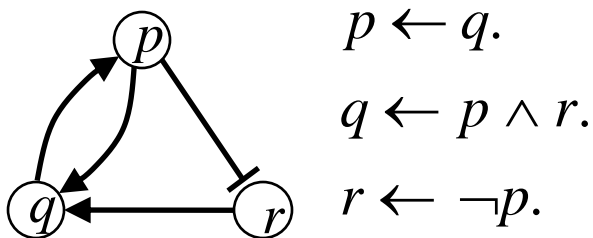


ダイナミック環境下における推論・学習



ブーリアンネットワークに関する推論

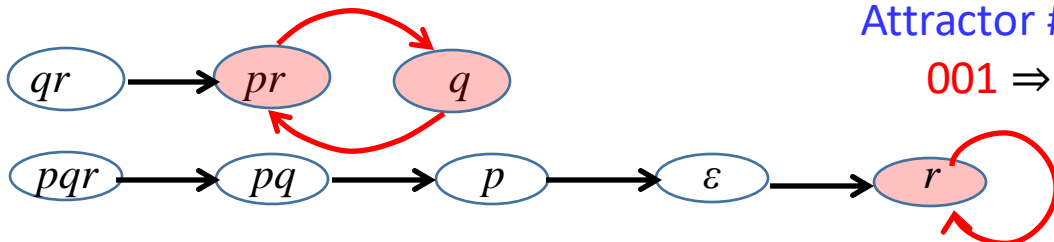
- ブーリアンネットワーク (Boolean Network): 遺伝子制御・シグナル伝達のような生体ネットワークやセルオートマトン等の複雑系、社会システム等を表現するための離散（ブール値）モデリング
- 正負フィードバックループを含むダイナミクスの解析



推論：論理プログラムの意味論
→ 軌跡とアトラクターの計算

Attractor #1: $(p,q,r) =$
 $101 \Rightarrow 010 \Rightarrow 101 \Rightarrow \dots$

Attractor #2: $(p,q,r) =$
 $001 \Rightarrow 001 \Rightarrow \dots$

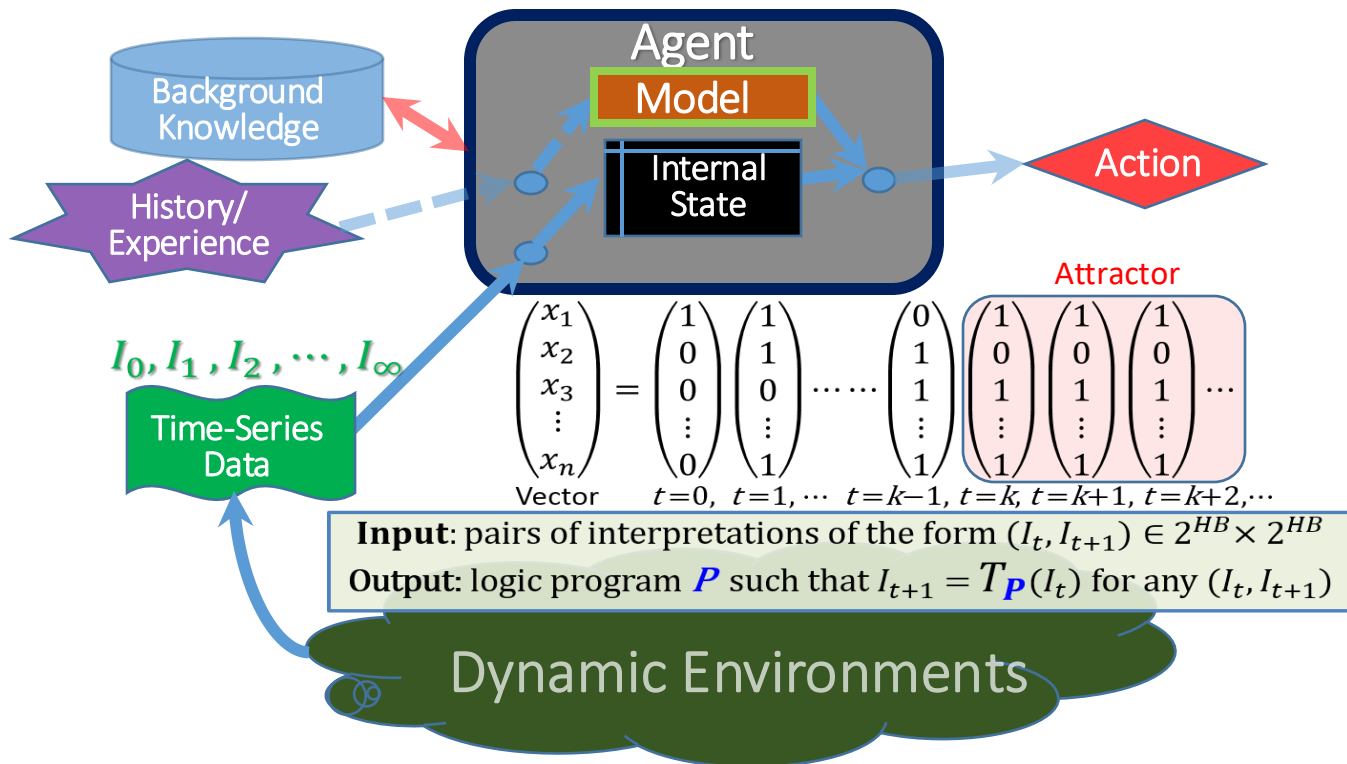


➤ Katsumi Inoue: Logic Programming for Boolean Networks. IJCAI 2011: 924-930.

➤ Katsumi Inoue, Chiaki Sakama: Oscillating Behavior of Logic Programs. *Correct Reasoning (Lifschitz Festschrift)*, 2012

解釈遷移からの学習

Learning from Interpretation Transition (LFIT)



Extensions:

- delays: Markov(k) system
- multi-valued variables^{*3}
- asynchronous update^{*3}
- nondeterministic^{*3} / probabilistic transitions^{*2}
- polynomial-time algorithm
- continuous domains
- neural semantics
- noise tolerance

Applications:

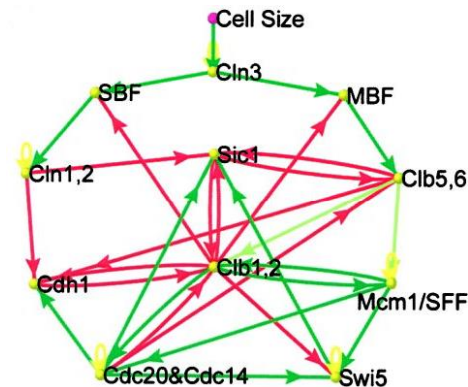
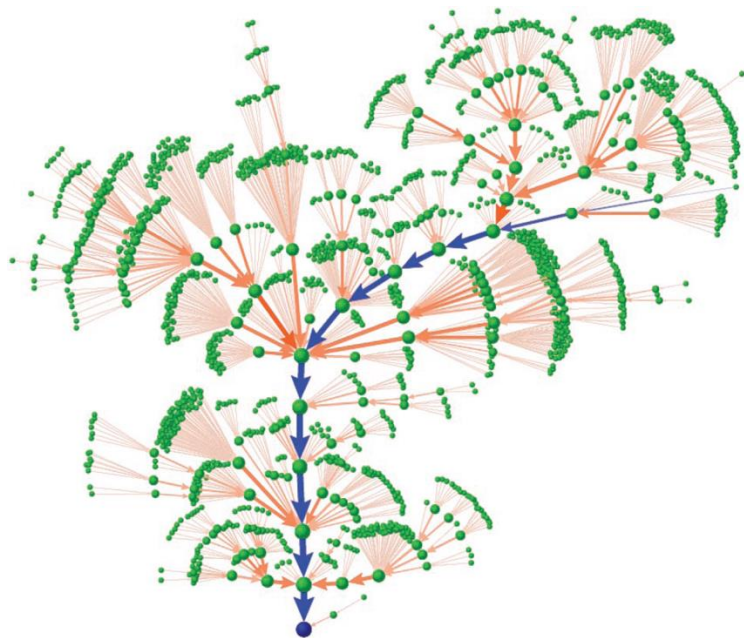
- biology (identification of gene regulatory networks)^{*1}
- DREAM challenges
- physics (cellular automata)^{*1}
- robots' actions in uncertain environments^{*2}
- learning agents' logics
- strategies in games
- intention recognition
- etc.

[*1] Inoue, K., Ribeiro, T., Sakama, C.: [Learning from interpretation transition](#). *Machine Learning*, **94**(1): 51-79 (2014)

[*2] Martínez, D., Alenyà, G., Ribeiro, T., Inoue, K., Torras, C.: [Relational Reinforcement Learning for Planning with Exogenous Effects](#). *J. Mach. Learn. Res.*, **18**: 78:1-78:44 (2017)

[*3] Ribeiro, T., Folschette, M., Magnin, M., Inoue, K.: [Learning any memory-less discrete semantics for dynamical systems represented by logic programs](#). *Machine Learning*, **111**(10): 3593-3670 (2022)

ロバストなブーリアンネットワークの学習



- Li, F., Long, T., Lu, Y., Ouyang, Q. and Tang, C.: The yeast cell-cycle network is robustly designed, *PNAS*, 101(14):4781–4786 (2004).
- Inoue, K., Ribeiro T., Sakama, C.: Learning from Interpretation Transition, *Machine Learning*, 94(1):51-79 (2014).
- Ribeiro, T., Inoue, K., Sakama, C.: A BDD-Based Algorithm for Learning from Interpretation Transition, ILP 2013, *LNAI*, Vol.8812, pp.47-63 (2014).
- 2^{12} の状態遷移の大部分は同一のアトラクターの引きこみ領域 (basin of attraction) に属する
- LFIT (ver. 2013) はこの状態遷移から 54 個の規則を 0.8 秒で学習した
- BDDを用いた改良アルゴリズム(2014)は同じ規則を0.18秒で学習

機械学習と記号推論の融合

Integrating Machine Learning and Symbolic Reasoning



記号推論の利点と弱点

機能


- 論理式で表現された前提知識（知識ベース）や仮定から論理的帰結を導く
- 論理式や数式により表現された制約式の集合を満たす解を計算する

メリット

- 無矛盾性・健全性・完全性といった計算の正当性を保証
- 推論結果は説明可能かつ解釈可能（XAIの基礎）
- 得られた結果を知識として蓄積し利用可能

デメリット

- 知識獲得ボトルネック：正しい知識を得ることの困難性
- スケーラビリティ・ノイズ耐性（ロバスト性）の欠如
- 常識推論：不完全情報からの推論が限定的

 **機械学習 (machine learning) との融合により克服する**
逆に、機械学習の信頼性を記号推論を導入することで向上させる

知識表現・推論と機械学習の融合

分離の河

記号世界 (離散)

❖ 知識表現・推論 (KR)

- 解釈可能性
- 説明可能性

Neurosymbolic AI

LLM/LRM

2つの世界の橋渡し

- ❖ 線形代数による論理プログラム計算
- ❖ 微分可能な論理推論・学習
- ❖ 機械学習への記号制約の導入

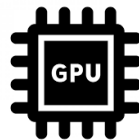
AI実応用

- ❖ 自動運転・物体認識
- ❖ 自然言語・VQA
- ❖ 工学・科学・ビジネス, 等

ベクトル空間 (連続)

❖ 機械学習 (ML)

- ロバスト性
- スケーラビリティ



GPUによる高速代数計算

進行中の研究テーマ

Neuro(-)symbolic AI

学習・推論に基づく汎用AI



Neuro-symbolic AI

- 知的行為における2つの認知的側面 (Leslie Valiant, 2003):
 - 経験からの学習 (神経回路網)
 - 学習したことからの推論 (記号処理)を統合するために、ニューラルAIと記号AIを融合するアプローチ
- Artur Garcez et al. (2019); Gary Marcus (2022); Francesca Rossi (2022), etc.
- 認知科学的には、脳の相補的アプローチと対応 (Daniel Kahneman, 2011):
システム1 (ニューラル/反射) ⇔ システム2 (記号処理/熟考)

各種のニューロシンボリックAI

- **ブラックボックス学習システムのための説明モデル**
 - ニューラルネットワークからのルール抽出
 - 機械学習モデルと同じ振る舞いをする論理回路の構築
- **ニューラル&記号のハイブリッドシステム**
 - ニューラルパターン認識→記号表現→記号推論による問題解決
 - 機械学習モデルの出力を記号推論により検証
 - ニューラルパターン認識を記号推論で強化 (学習候補の制約等)
- **記号知識のベクトル空間への埋め込み(embedding)**
 - 連続空間における推論：定理証明, 論理プログラミング, アブダクション
 - 知識グラフ(Knowledge Graph)埋め込み
 - プログラム合成, 微分可能プログラミング

➤ C.f. 6 patterns of neuro-symbolic systems by Henry Kautz (2020)

機械学習と記号推論の融合に向けて 現在取り組んでいる課題

A) 代数的手法に基づく論理表現・推論 (*)

1. 線形代数(linear algebra)を用いた論理プログラミング
2. ↳ 微分可能(differentiable)論理プログラミング

↓ **代数的手法の適用**

B) 制約を取り入れたAI・機械学習システム

1. 自動運転・物体認識等への制約の導入
2. 制約を考慮した生成AI



信頼される AI システム
Trusted quality AI systems



(*) Katsumi Inoue: [Algebraic Connection Between Logic Programming and Machine Learning](#) (Keynote), FLOPS 2024

代数的手法に基づく論理表現・推論 (1)

—線形代数を用いた論理プログラミング

- 共通する原理 (Sakama, Inoue & Sato, 2017;2020):
 - 論理表現(encoding): ベクトル・行列・テンソルを用いる
 - 計算(computation): これらの要素に対して線形代数演算を適用する

- P : (論理)プログラム・制約 \rightarrow 行列 M_P
- I : (真偽)値割り当て/解釈(interpretation) \rightarrow ベクトル v_I
- $J = T_P(I) = \{h \in HB \mid (h \leftarrow b_1 \ \&\dots\ \& b_m) \in P \text{ and } \{b_1, \dots, b_m\} \subseteq I\}$: 1ステップ推論
- このとき、 J は行列・ベクトル積と閾値関数 θ を用いて次の式で計算可能:

$$v_J = \theta(M_P \cdot v_I)$$

- 期待される効果:
 - 行列のスパーシティを利用した高速計算 (Nguyen, Inoue & Sakama, 2022)
 - GPU計算による並列性・スケーラビリティ + 部分計算による指数的高速化

代数的手法に基づく論理表現・推論 (2)

—連続空間上で微分可能な推論・学習

- 共通する原理 (Sato & Kojima, 2019):

- 損失関数(loss function) J の設定
- テンソルパラメータ \mathbf{x} をもつ J の連続最小化問題として定式化
- 確率的勾配降下法 (SDG)/Newton法により J を最小とする \mathbf{x} を計算
- $J(\mathbf{x}) = 0$ ならば \mathbf{x} は解 (真偽値割り当て・安定モデル・支持モデル等)
- \mathbf{x} を2値テンソルに閾値化(thresholding) (非微分可能部分)

$$\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}}$$

- 期待される効果:

- 連続性によるロバスト性・ノイズ耐性
- マルチコア/GPU並列計算によるスケーラビリティ
- ニューラル系・機械学習との接続容易性

ベクトル/テンソル空間上に実現した各種の論理推論

一階述語演繹
(Sato, TPLP 2017)

一階述語アブダクション (Sato, Inoue & Sakama, IJCAI 2018)

論理プログラミング(LP)不動点計算 (Sakama, Inoue & Sato, KSEM 2017; AMAI 2021)

スパース法による LP 計算 (Nguyen, Inoue & Sakama, ICLP 2021)

ASP (supported models) (Sato, Inoue & Sakama, ICAART 2020)

LP アブダクション (Nguyen, Inoue & Sakama, ICTAI 2021; PADL 2023; ICTAI 2024)

微分可能 ASP (supported models) (Takemura & Inoue, LPNMR 2022; ECAI 2024)

微分可能 ASP (stable models) (Sato, Takemura & Inoue, arXiv 2024)

SAT (MatSat) (Sato & Kojima, PoS 2021)

ブーリアンネットワーク学習 (Sato & Kojima, KR 2021)

微分可能 LFIT (transformer-based) (Phua & Inoue, ILP 2019; ILP 2021; NeSy 2024)

微分可能 LFIT (matrix learning) (Gao, Wang, Cao & Inoue, MLJ 2022)

一階述語 LP 学習 (Gao, Inoue, Cao & Wang, IJCAI 2022; AIJ 2024)

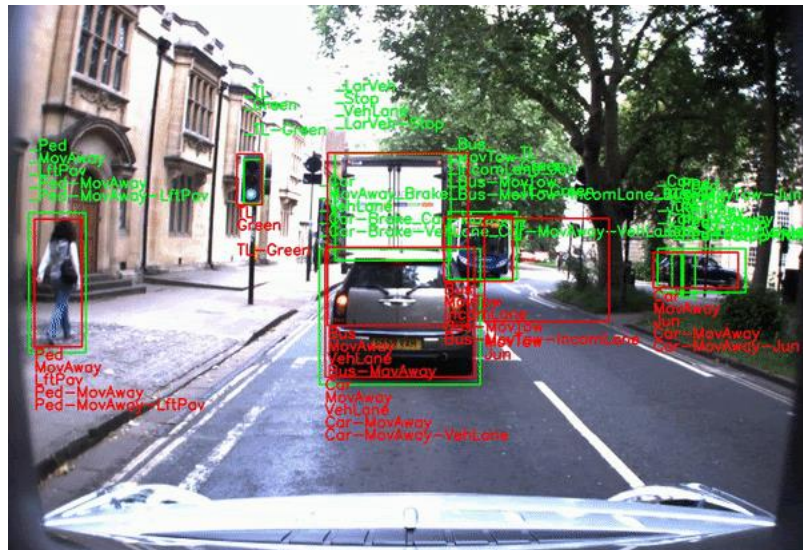
DNF 学習 (Sato & Inoue, MLJ 2023)

時系列データからのプログラム学習 (Gao, Inoue, Cao, Wang & Yang, ICLR 2025)



ROAD-R Challenge for NeurIPS 2023

- 深層学習による物体検知と自動運転への応用
- マルチラベル物体認識
 - 物体 × アクション × ロケーションの認識
 - **ROAD: ROAD Awareness Dataset**
- 論理的ハード制約（必要条件）の導入
 - 車線には駐車できない
 - 信号は同時に赤と青にはならない
 - エージェントが横断していれば、歩行者か自転車に乗っている



<https://sites.google.com/view/road-r/dataset>

- **ROAD-R: ROAD with Requirements [1]**
- 手法：事前訓練済み物体検知アルゴリズムを拡張 [2], Partial Weighted MaxSAT を適用
- [NIIチーム結果: Task 2 \(制約付き認識\) 優勝, Task 1 \(物体認識のみ\) 3位](#)
- 今後、他の領域における制約付き動画理解や制約付き画像・動画生成AIに展開

[1] Eleonora Giunchiglia, et al.: ROAD-R: the autonomous driving dataset with logical requirements. *Machine Learning*, 112 (2022)

[2] Moriyama, S., Watanabe, K., Inoue, K., Takemura, A.: MOD-CL: Multi-label Object Detection with Constrained Loss. arXiv (2024)

展望：記号推論による知性

- 機械学習と知識表現・推論の結合はさらに進む
 - 推論の本質はさほど変わっていない
 - クラスNPを超える記号推論の活用
 - 常識推論の理解、LLMとの結合
 - 演繹以外の推論を機械学習・生成AIと結合
 - ノイズがある環境におけるアブダクション
 - 連続領域が扱えるロバストな記号推論 \leftrightarrow ファジイ論理？
 - 認識・学習・推論という一連の知能を共通の土台で扱う
 - 意識・理解・精神性等の問題をどう扱うか

多くの研究者や大切な人々と出会い
先生方・先輩・学生諸君に励まされ
職場や周りの方に大変お世話になり
暖かく支援していただいたおかげで
研究を続けてくることができました

みなさま、ありがとうございました

2025年3月27日

井上 克巳