

発表要旨

これまでの研究業績

Fully Automated Cyclic Planning for Large-Scale Manufacturing Domains ≈ 1. **ICAPS14**.

Solving Large-Scale Planning Problems by Decomposition and Macro Generation ≈ 1. **ICAPS15**.

Tiebreaking Strategies for A* Search: How to Explore the Final Frontier ≈ 1. **AAAI16**. (**JSAI** 学生奨励賞)

Tie-Breaking Strategies for Cost-Optimal Best First Search ≈ 1. **JAIR** **58 (2017): 67-121**.

Exploration Among and Within Plateaus in Greedy Best-First Search ≈ 1. **ICAPS17**.

Efficient Optimal Search under Expensive Edge Cost Computation ≈ 2. **IJCAI17**.

Classical Planning in Deep Latent Space: From Unlabeled Images to PDDL (and back) ≈ 1. **KEPS17**.

今後の研究計画、研究成果の産業応用への抱負

≈ 1 Masataro Asai, Alex Fukunaga

≈ 2 Masataro Asai, Akihiro Kishimoto, Adi Botea, Radu Marinescu, Elizabeth Daly M, and Spyros Kotoulas

1 背景 - AI プランニング



1.1 誰?



And let me introduce these robots. The guy in the left is Astro boy.

1.1.1 誰?

鉄腕アトム



1.1.2 誰?

鉄腕アトム



見て,聞いて,喋って,行動する
自律行動 *Autonomous*



As you know, he is a famous manga superhero invented by Tezuka Osamu in 50s,

1.1.3 誰?

T-52 援竜 @北九州消防局
レスキュー・ロボット

鉄腕アトム



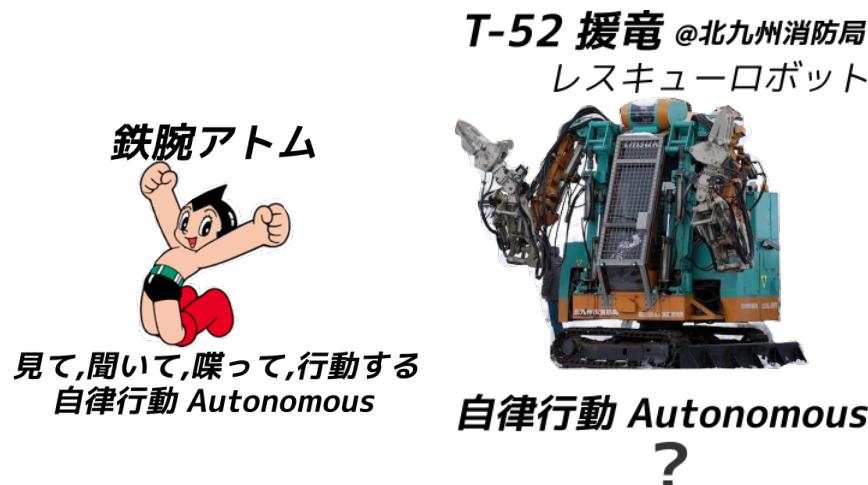
見て,聞いて,喋って,行動する
自律行動 *Autonomous*



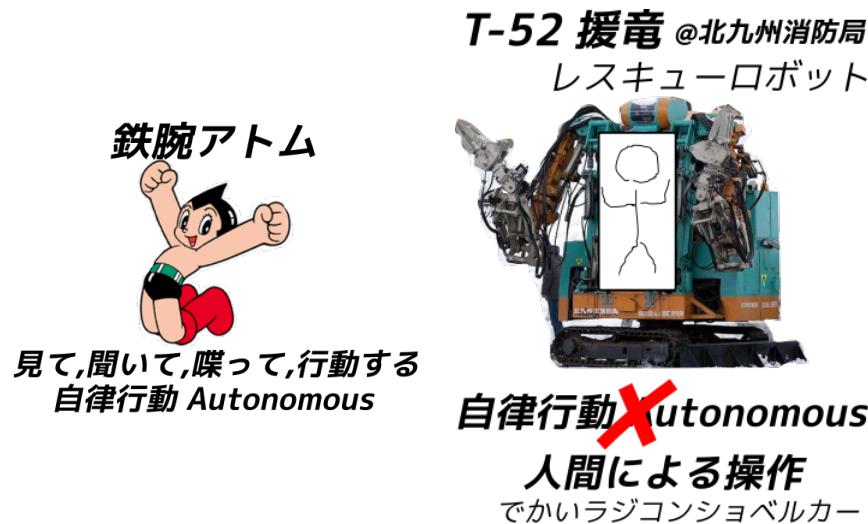
and he can think, hear, speak, act. he also has emotions.

In contrast, the guy in the right is a real robot that is actually in use @ fukuoka prefecture for the rescue purpose.. His name is T-52 Enryu, developed by a Japanese company Temzak. He is huge and powerful – about 4 meters in height and can carry things which is as heavy as 500kg. Well, so, in a sense, he is also a superhero in the real disastrous situation.

1.1.4 誰?

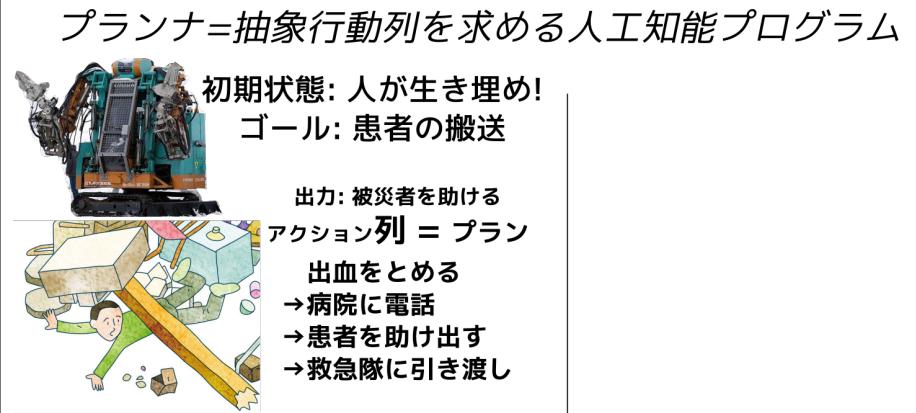


1.1.5 誰?



But does he have feelings or can he think? Can he even move around by his own?

1.2 自律行動のための自動プランナ (= モータ制御)



No. It requires full human intervention — it is indeed operated by a driver who gets in or by a remote control. It is more like a super-sophisticated shovel car.

研究テーマのプランニングは、ロボットに、人間の助けを借りず、いかに自律して行動させるかを扱います。これをモデル化したプランニング問題は、具体的な行動の列を求める組合せ最適化問題です。

プランニング問題のタスクは、センサーから初期状態とゴールを受け取って、被災者を助ける正しい手順を出力することです。

たとえば、この図では男性が瓦礫に埋まって助けを求めています。プランニング機能のあるロボットは、コレに対して「男性を助けよ」という大まかな指示を受けています。

1.3 自律行動のための自動プランナ(≠ モータ制御)

プランナ=抽象行動列を求める人工知能プログラム



指示の内容には、図のように初期状態とゴール、許可された行動のリストが入っています。ロボットは、自動プランニングにより、人間の代わりに適切な行動を組み立てて、ゴールを自動で達成します。

1.4 自律行動のための自動プランナ(≠ モータ制御)

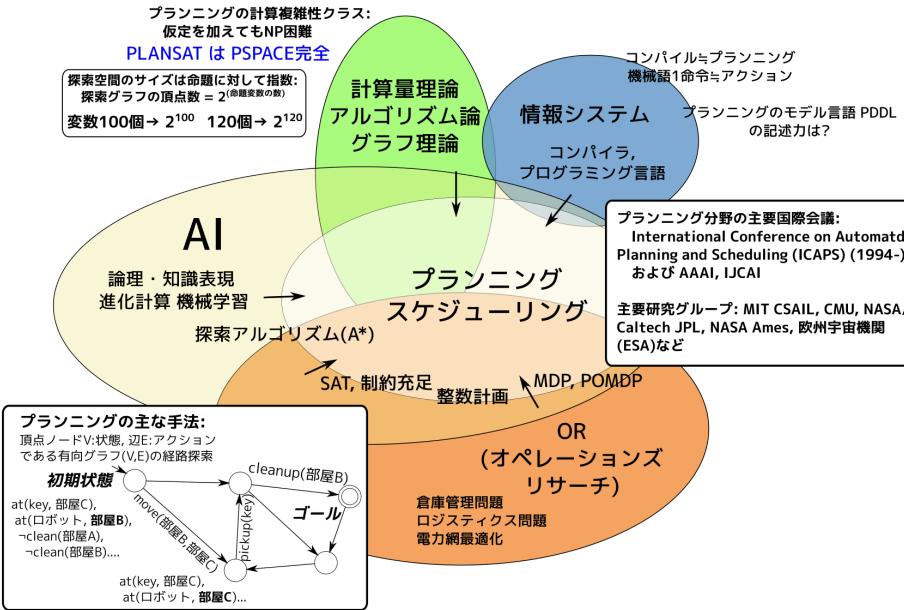
プランナ=抽象行動列を求める人工知能プログラム



プランニングは汎用な枠組みなので、災害救助以外にも様々な問題に適用することができます。現実の応用例では「宇宙探査機運行問題」や「企業ネットワーク脆弱性問題」も表現できます。

このように、プランニングは、難しい問題を汎用性を失わずに解くことを目指します。

1.5 プランニング(自動行動計画)分野の位置づけ



プランニング分野は、人工知能の専門分野という位置づけで、隣接するオペレーションズ・リサーチやアルゴリズム論などの分野の技術を利用してしています。特に、プランニング問題を解くのにはグラフ探索の技術が用いられます。

1.6 業績1: 査読付き国際学会 ICAPS14 (採択率33%)

セル生産方式: 工場生産方式の一種(対義語: ライン生産)

ロボットアームなどを含む小規模セルが複数協調して製造

目的: 並列の製造プランを自動計画 + 製作時間を最小化

一台の組み立て問題を解くのはある程度複雑でも可能

研究の対象

× アクチュエータの角度制御 (例: θ: 0deg → 30deg, 1deg/sec)

◎ 行為の計画 (例: 部品2をアーム1で塗装, ネジ3を締結 etc.)

問題点: 扱う対象の数が多くなる場合(製品が4個程度以上の時)

最適解を返すソルバで直接解く → 探索空間が広大で探索不可

1つの製造手順を複数つなげて後処理最適化 → 生産時間が長い (makespan)

手法: 問題の繰り返し構造を「工場, セル生産, 製品」など前提なしに

任意のプランニング問題の論理構造から自動検出

検出した繰り返し構造から効率よい(生産時間の短い)プランを生成

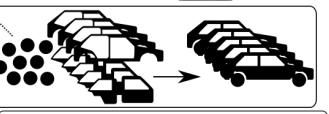
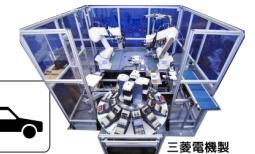
結果: 1000台以上の製品を組み立てる問題を解けた

+ 生産時間を最大3割短縮(後処理最適化と比較)

新規性: 繰り返し構造を自動で抽出するというアイディア+抽出技術

インパクト: 大きな高速化、プランニングの適用規模を広げた

株式会社IHIと共同研究(システムの先行評価として)



探索空間のサイズは指數: $2^{\text{命題変数数} \times \text{数}}$
変数100個 → $2^{100} \rightarrow 120 \rightarrow 2^{120}$
問題が少しきくなるだけで百万倍速くなる

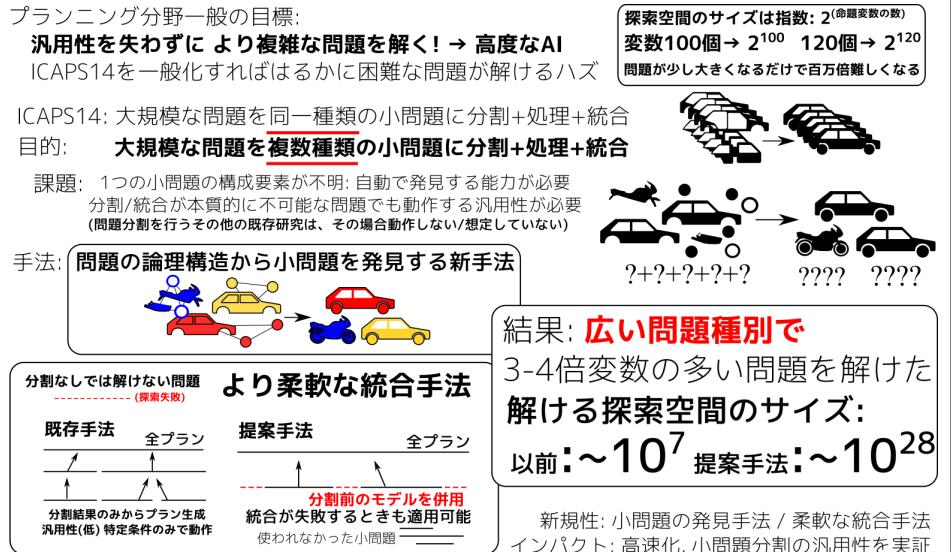
→ しかも探索空間はその指數で増加
解ける探索空間のサイズ:

以前: $10^6 \rightarrow$ 提案手法: 10^{274}

研究業績に移ります。ここでは、大規模なプランニング問題を解くために問題設定・ドメインによらず汎用に繰り返し構造を抽出する方法を開発しました。元の問題を繰り返し一周分の小問題に分割して解くことで、高速化と3割の生産時間短縮を達成しました。

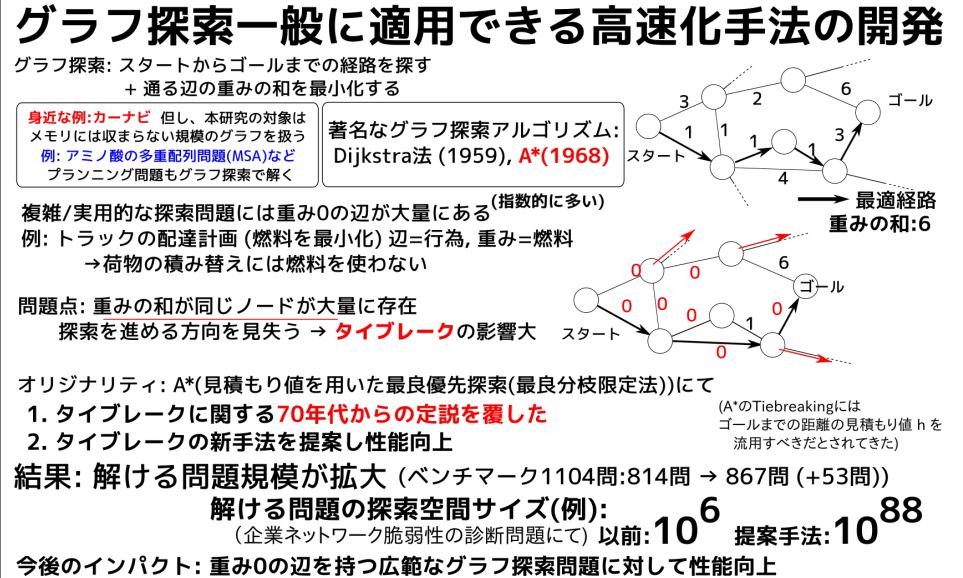
大切なのは汎用性です。生産問題に限らず、掃除にも宇宙船にも同じ実行バイナリが使えます。

1.7 業績2: 査読付き国際学会 ICAPS15(採択率33%)



続いて二、三本目の業績は、先ほどの手法で得られるのは1種類の小問題だけでしたが、これを複数種類の小問題に拡張しました。結果、より様々な問題で高速化を達成しました。ここまで汎用に小問題分割を適用した研究は、分野では始めてです。

1.8 業績3: 査読付き国際学会 AAAI16(採択率26%)



最後に、申請後に行った研究が、難関国際学会 AAAI に採択されました。研究内容は、コスト 0 の辺を含むグラフを扱うグラフ探索アルゴリズム一般に適用できる内容で、非常に大きなインパクトを持つことが考えられます。

1.9 [

業績4: 査読付き論文誌 JAIR (採択率 12%)] 業績4: 査読付き論文誌 JAIR (採択率 12%)

スキップ予定

AAAI16の一般化: タイブレーク≡非最適探索

問題点: 重みの和が同じノードが大量に存在 (ゼロコスト問題)

探索を進める方向を見失う → タイブレークの影響大

ノードの下界関数 h の種類

許容的: 常に真のコストより低く見積もる → 最適解

非許容的: 非最適解を見つてしまふ ← 従来は使われない

Unit cost 関数: 辺コストを全て1で代用, 非許容的関数の例

主な発見: タイブレーク とは 許容的関数 h で作ったプラトーの中で非最適探索を行うことである

新規性: 最適探索で非許容的関数を効果的に使う方法を初めて示した

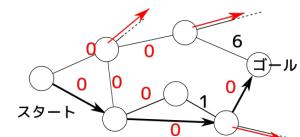
性能評価:

ゼロコスト問題のタイブレークに

非許容的 Unit cost 関数を使う

→ 指数的に高速化 (844問 → 906問)

→ AAAI16と同時使用でさらに高速 (→ 911問)



理論的成果:

無限グラフでのA*の完全性を

様々なタイブレーク法ごとに証明

→ 非最適探索の理論的結果を援用

非最適探索用の全技術を最適探索に援用可能
(従来は非互換と考えられてきた)
→ 今後のさらなる発展が期待できる

1.10 [

業績5: 査読付き論文誌 JAIR (採択率 12%)] 業績5: 査読付き国際学会 ICAPS17 (採択率 33%)

スキップ予定

非最適探索の 均一化手法 の理解と改善

非最適探索: GBFS (貪欲最良優先探索) + 非許容的 下界関数 h

(exploration/diversification)

下界関数 h を常に信頼して(貪欲)探索を集中させる(exploitation)

→ 非許容的 h は常に正しいとは限らない → h を時折無視して探索を均一化 → 性能向上

問題点: アドホック手法が多い

→ どう組み合わせればよいか解らない

目的: → クリーンに評価しなおす



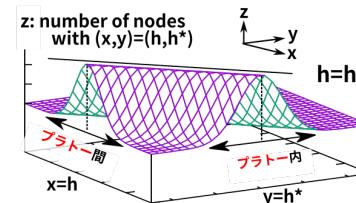
発見1. 2つの直行した均一化方法がある。

理由: 真の値 h^* からの h の誤差は二次元的

→ プラトー間均一化と プラトー内均一化

→ 同じ均一化手法を二通りに利用出来る

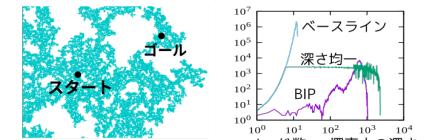
→ 直行だから同時利用してOK! → 性能向上



発見2. ボンドパーコレーション(BIP)

フラクタルを使った均一化手法

→ 性能向上の理由: 探索木の幅を削減



制限時間内で解けた問題数: 192 → 237.7
参考: [Xie et. al. 2014] 深さ均一化 223.9

1. 未知の均一化アルゴリズムに汎用に適用できる法則
2. フラクタルを用いてより良い探索が出来る可能性

1.11 業績6: 査読付き国際学会 IJCAI17 (採択率 25%)

辺コスト動的計算が必要なグラフ探索の高速化

(IBM Research Ireland での研究)

例1: 辺コスト(都市間距離)が未知のTSP

都市数Nに対し $O(N^2)$ 回経路探索ソルバを呼び出す

例2: MWRP(複数ワーカー乗り換え問題):

[地図上の全場所] $^2 \times$ [全出発時刻] 個の二点間乗り換え

→ 全ての辺コストを求めるのは現実的でない

→ コストを求めた後の問題自体もNP困難

在宅訪問診療のモデル問題

医者が患者を予約時間に訪問する最適乗り換え計画



問題点: 一回の辺コストを求める計算(外部ソルバ)が重い

→ 全コストを求めるのが現実的でない

→ A*探索で必要になった辺コスト c_a のみを動的に計算

→ それでもなお実行時間の90%を辺コスト計算に使用

解決法: c_a の下界 c_h で代用 必要になるまで c_a 計算を遅延

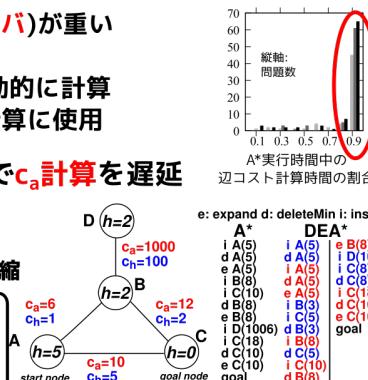
c_h : 外部ソルバの下界閾値数, c_a より高速に計算可能

同じノードを c_h/c_a で二回展開・キューに挿入

→多くの c_a を計算せずにすむ → 探索時間を最大5倍短縮

多くの実問題は階層構造を持つ

既存外部ソルバを用いて汎用に適用可能



2 産業技術総合研究所で行いたい研究

Neural-Symbolic 複合システムによる

る

次世代AIシステム

の研究

2.1 Q. いまはやりのDeep Learningとの違いは?

A. レイヤが違う

機械学習・Neural Networks ==

関数近似

for 認識・反射

- 入力は **Subsymbolic** (連続値)

画像、音声、非構造化テキスト:

- 感覚的知能:

反応、直後の行動の決定
パブロフの犬: 飼料を認知 → よだれ

自動運転: 赤信号、人 → 止まる。

翻訳: 文章 → 文章

囲碁局面の評価関数: 局面 →

勝率 効率よく 1-to-1

mapping

単純作業

- AlphaGo = Subsymbolic (DLNNによる評価関数) + Symbolic (MCTSによる探索)

推論・探索

for プランニング・ゲーム・定理証明

- 入出力は **Symbolic**

論理オブジェクトルール

- 論理・推論による知能:

未来に渡る戦略の決定
(戦略 = 行動の列や木)

レスキューロボ: ゴール = 被災者生存

証明器: ゴール = QED

コンパイラ: 命令列の生成

囲碁、将棋: ゴール = 勝利 順序制約+複雑な作業

2.2 既存の有名システム

AlphaGo = Subsymbolic (NN による評価関数) + Symbolic (MCTS による探索)

- ただし ドメイン依存 – 囲碁に特化, "マス目"や"石"といった概念をハードコード
- 膨大な棋譜が必要 — 運用データがない環境 (e.g. 火星) には適用不能
- 人って模範解答がないと行動できませんか? 真の自律機械は前例無しでも行動可能

DQN = Subsymbolic (DLNN) + 強化学習 (DLNN)

様々な Atari Game につかえる汎用フレームワーク (Invader, Packman ...) だが

- RL の Acting: 学習した policy に従って greedy に行動
- Atari ゲームは 脊髄反射で生き残ることが可能 → 複雑な論理思考はいらない!

2.3 記号的 AI による論理推論の重要性

f Conclusions

Y LeCun

- Deep Learning is enabling a new wave of applications
- ▶ Today: Image recognition, video understanding: [vision now works](#)
- ▶ Today: Better speech recognition: [speech recognition now works](#)
- ▶ Soon: Better language understanding, dialog, and translation

Deep Learning and Convolutional Nets are being widely deployed

Today: Image understanding at Facebook, Google, Twitter, Microsoft....
Video analysis, medical diagnosis, autonomous vehicles, robotics, etc.

We need hardware (and software) for embedded applications

- ▶ For smart cameras, mobile devices, cars, robots, toys....

But we are still far from building truly intelligent machines

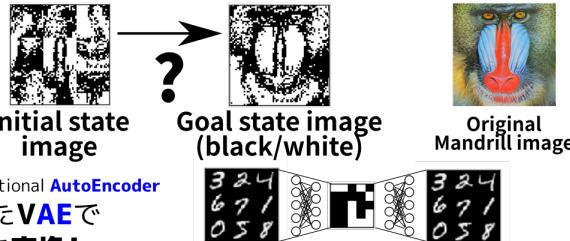
- ▶ We need to integrate [reasoning](#) with deep learning
- ▶ We need a good architecture for "episodic" (short-term) memory.
- ▶ We need to find good principles for [unsupervised learning](#)

2.4 業績7: 査読付きワークショップ KEPS (採択率 60%)

深層学習(非記号型AI)と行動計画(記号型AI)の融合

画像で示された問題を

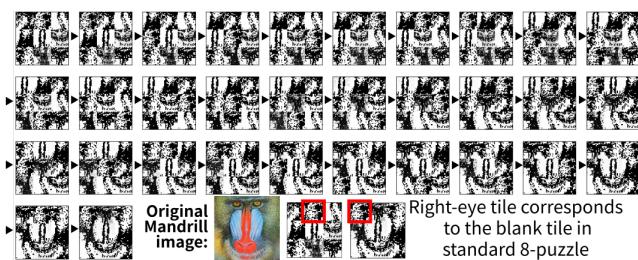
「パネル」「9マス」「動く」などの情報は一切与えられない問題を解いた例も与えられない



ノードの数を減らしてから元のサイズに戻すNN
入力画像と出力画像が一致するように学習
かつ、真ん中の層は0/1値: 何らかの命題の真偽値に対応

結果を画像に戻して
プランを返却する
システム

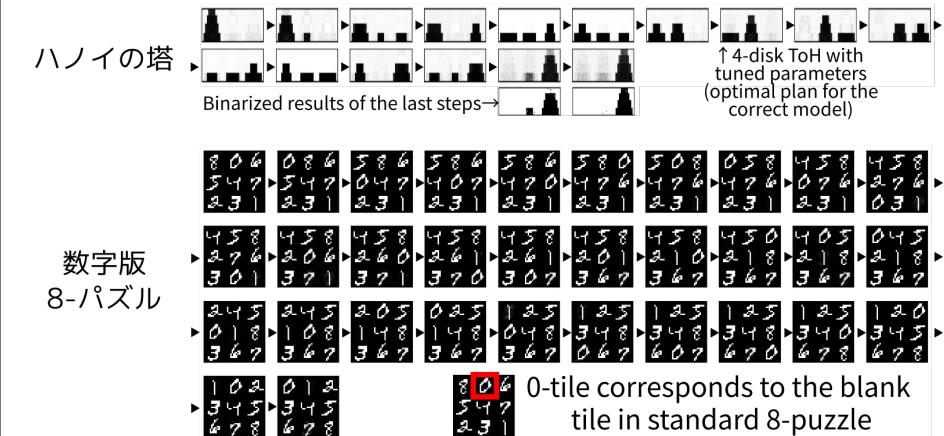
LatPlan



2.4.1 研究業績7: 査読付きワークショップ KEPS

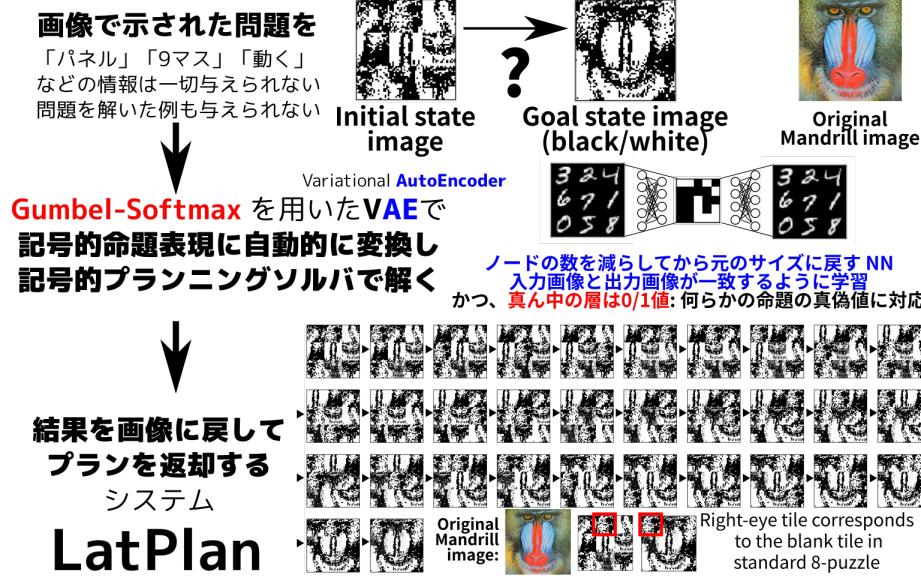
LatPlan: 同じシステムで全く異なる問題を解ける

問題ごとにニューラルネットの学習は必要



2.4.2 業績 7: 査読付きワークショップ KEPS (採択率 60%)

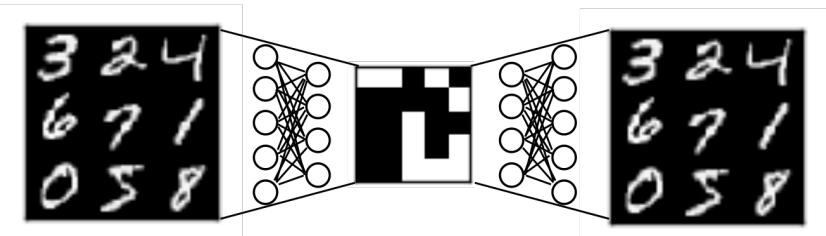
深層学習(非記号型AI)と行動計画(記号型AI)の融合



2.4.3 Gumbel-Softmax

AutoEncoder(AE): 入力と出力が一致するように学習するNN
中間層: 記号的システムで扱えない実数値

Gumbel-Softmax AE: 1-hot カテゴリカル分布を近似するAE
中間層: 0/1値, 何らかの命題の真偽値に対応させて記号的システムで使える!



2.4.4 研究業績 7: 査読付きワークショップ KEPS

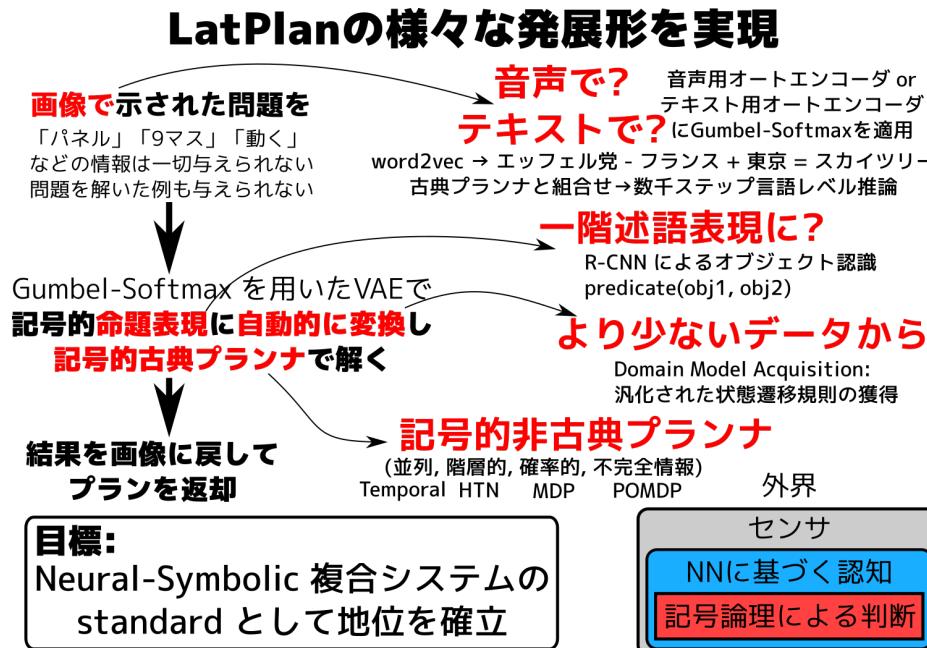
強化学習とは異なり、優れた理論的性質

アルゴリズムの完全性: 解が存在するときには必ず解を発見する
解の最適性: 理論的性質のわかっている下界関数を使うことで保証

	入力・認識	意思決定
DQN	画像・ニューラル (事前知識なし)	Greedy+NN強化学習 =反射的エージェント
AlphaGo	ハードコード ("石", "マス目")	記号的 MCTS+NN強化学習 完全性 vs. イ・セドル 白74手 「読み違え」 学習結果に理論保証なし → 非最適解 ($t \rightarrow \infty$ での収束保証のみ)
LatPlan	画像・ニューラル (事前知識なし)	A*+許容的下界関数 (完全性+解の最適性)

認識の誤り(NNに起因) はあれど 決断の誤りは無い
ニューラル認識と記号的意思決定(理論保証付き)を持つシステム

3 今後の研究計画



4 研究成果の産業応用

記号的推論により様々な分野にブレークスルーをもたらす

ピックアップロボ以上の産業用

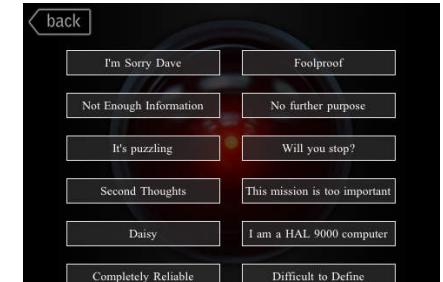
ロボ

— 画像 → 論理表現 → 記号的推論 → 行動決定

— 動的な環境や目標に自動で対応できる

人工無能以上のチャットボット

— テキスト → 論理表現 → 記号的推論 → 意図推定 → 返答



5 まとめ

1. 難関国際会議 (33%) Fully Automated Cyclic Planning for Large-Scale Manufacturing Domains. In ICAPS2014.
 - (a) 任意の問題から 1種類の繰り返し構造を自動で検出
 - (b) 工場での製造スケジューリング (x1000 高速化, 探索空間 $10^6 \rightarrow 10^{274}$)
2. 難関国際会議 (33%) Solving Large-Scale Planning Problems by Decomposition and Macro Generation. In ICAPS2015.
 - (a) 複数の繰り返し構造をより柔軟・汎用に組み合わせる手法
 - (b) ベンチマークセット全体で高速化 (x3-4 高速化, 探索空間 $10^7 \rightarrow 10^{28}$)
3. 難関国際会議 (26%) Tiebreaking Strategies for A* Search: How to Explore the Final Frontier. In AAAI-2016. (JSAI 学生奨励賞)
 - (a) コストゼロの辺がグラフ探索に引き起こす問題を解決 (探索空間 $10^6 \rightarrow 10^{88}$)

以上です、ありがとうございました。

4. 難関論文誌 (12%) Tie-Breaking Strategies for Cost-Optimal Best First Search. *Journal of Artificial Intelligence Research* 58 (2017): 67-121.
 (a) (3.) に加え タイブレーキングと 非最適コスト探索の関連性を指摘、さらに性能向上
5. 難関国際会議 (33%) Exploration Among and Within Plateaus in Greedy Best-First Search. In ICAPS2017.
 (a) 非最適コスト探索をフラクタルを用いて改善
 (b) プラトー内均一化とプラトー間均一化の直交性を実証
6. 難関国際会議 (25%) Efficient Optimal Search under Expensive Edge Cost Computation. In IJCAI-2017.
 (a) 辺コストの動的計算が必要な問題に対して高速な最適アルゴリズム DEA*
7. 国際ワークショップ (60%) Classical Planning in Deep Latent Space: From Unlabeled Images to PDDL (and back). Knowledge Engineering for Planning and Scheduling (KEPS) Workshop
 (a) 画像から命題を自動生成して記号的 AI で組合せ最適化問題を解き、画像で出力するシステム

6 付録

6.1 既存システムとの違い (追加)

NN で直接問題を解くシステム

TSP [Hopfield and Tank, 1985], NeuroSolver [Biesczad and Pagurek, 1998]

— NN で解くが、入力はシンボリック (それぞれのニューロンが人の与えた状態変数に対応)

— 完全性、最適性などの補償なし

NN を認識でなく探索の枝刈りの中で使うシステム

AlphaGo [Sievers 16], [Arfaee et al., 2011], [Satzger and Kramer, 2013]

— LatPlan は NN を探索の外で使う

6.2 Learning from Observation との違い

主にロボットの経路探索 (ローレベル制御) [Argall et al., 2009]

ボードゲームの学習だが「マス目」など強い仮定 [Barbu et al., 2010; Kaiser, 2012; Kirk and Laird, 2016]

Action Segmentation problem がある

— 「映像の観察」を中心とするので、いつアクションが始まる/終わるのか解らない

— LatPlan には関係なし