RIKEN AIP-NEC Collaboration Center
理研AIP-NEC連携センター

理研AIP-NEC連携センター設立の狙い

AIの社会実装を加速させ、安全・安心で効率的な社会を実現

実現に向けた課題 → **課題解決に向けた“基盤技術”を開発**

- AIの社会実装加速に大きな壁となる大量学習データの不要化 → 活動テーマ１、２
- AI活用による自動化が進み、複数AI間での競合がある中での最適化 → 活動テーマ３
- 社会実装の加速に向けたAIの社会受容性の考慮 → 活動テーマ４
- AIのチューニング（ハイパーパラメータ最適化）手間削減 → 活動テーマ５

活動５テーマと活動期間・内容サマリ

- **3テーマで活動を開始し、社会動向やAIの技術進展動向に応じてテーマを新設・変更しながら成果を創出**

活動テーマ

	～2019	2020	2021
1．少量の学習データで高精度を実現する学習技術の高度化	異常事象など一部種類の現場データ無い場合に適応する学習手法	真のラベルよりも低コストで入手可能な弱ラベルを用いた学習手法の改善	
	（ACML2019,WACV2020採択）		（ICML2021採択）
2．未知状況での意思決定を支援する学習／AI技術の高度化	論理推論のロバスト性を強化（観測との矛盾を高速検知）	サイバー攻撃手口推定の実問題で有効性確認	卒業
	（産総研含めた三者連携、NeurIPS2019採択）		
3．複数AI間の調整にかかわる学習技術の理論解析	AI間の交渉戦略理論：未知制約付きベイズ最適化高速手法を考案	製造・物流で広く使われるスケジュール調整問題をターゲットにした最適化手法	
	（産総研含めた三者連携、SIP）		（今後、学会発表予定）
4．実環境を想定したAXのリスク評価・対策	新規立上	顔認証の運用を加味したAX攻撃手法の提案・対策手法を考案	卒業
		（Biosig2021, IJCB2021）	
5．汎用的、パラメータフリーなブラックボックス最適化技術の開発	新規立上	理論的に収束性を示す新手法を考案	
			（今後、学会発表予定）

テーマ１ 少量の学習データで高精度を実現する学習技術の高度化

- **産業応用上重要となる学習データ獲得困難な事例（例：異常検知、エキスパートの代替、新規事象への迅速対応、特殊センサの認識）に対応**
- **ターゲットとしては画像認識のアプリケーションを想定するが、本連携活動では、特定アプリに特化しない、汎用的な機械学習技術の構築を目指す**

- **3タイプの学習データ獲得困難事例**

①現場の多様なニーズ（新たな事象）への対応が必要な場合

- 毎回、大量の学習データを作成していては、時間的・人的コストがかかりすぎる
- ニーズへの迅速な対応ができない

②データの発生頻度が少ない場合

- 異常データなどは、データ収集に時間がかかる
  例）1日1件発生しても、5,000件集まるまでに13年以上かかる

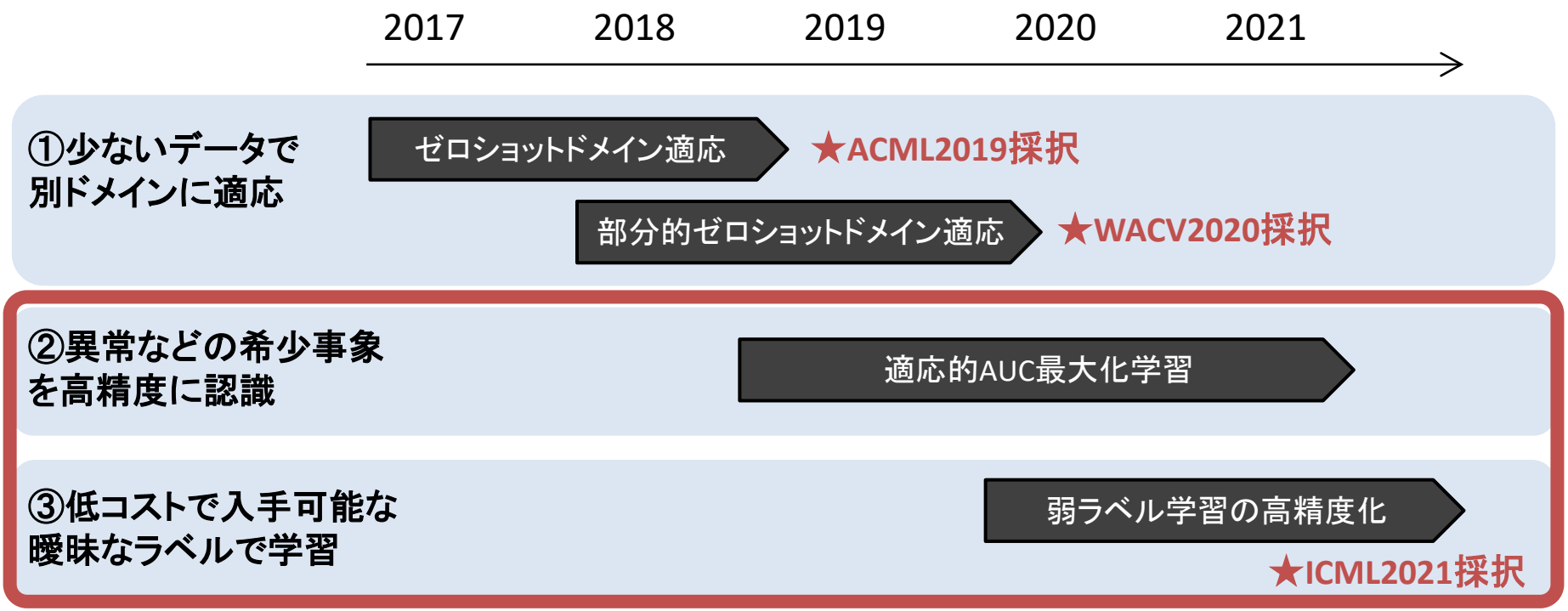
③正解付けコストが非常に高い場合

- 医療データなどは、専門家でないと正解がつけられない

テーマ１：成果概要

活動の振り返り

●深層学習の実用化に向けた問題を設定して、それぞれ基盤技術を開発

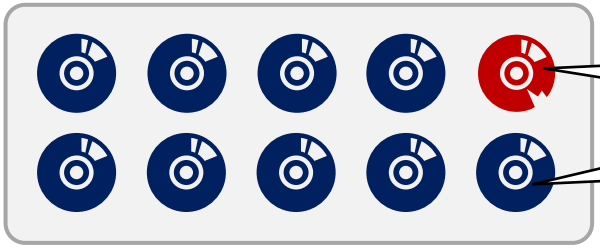


テーマ1：適応的AUC最大化学習

背景：異常などの希少事象のデータは、正常データに比べて非常に少ない

⇒ クラス間でデータ数が大きく異なる状況（クラス不均衡）

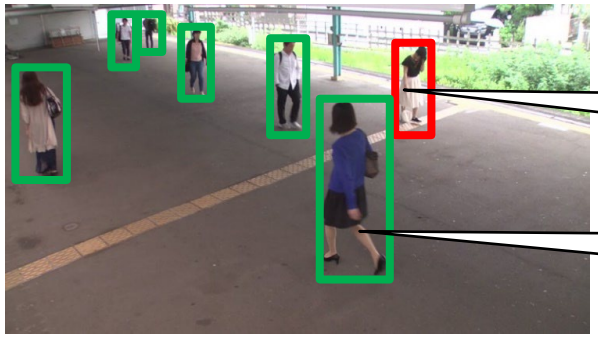
工業製品の検査検品



不良品（異常）の
データは**非常に少ない**

正常な製品の
データは**非常に多い**

監視カメラ映像からの
ふらつき歩行検知



ふらつき歩行（異常）の
データは**非常に少ない**

通常歩行（正常）の
データは**非常に多い**

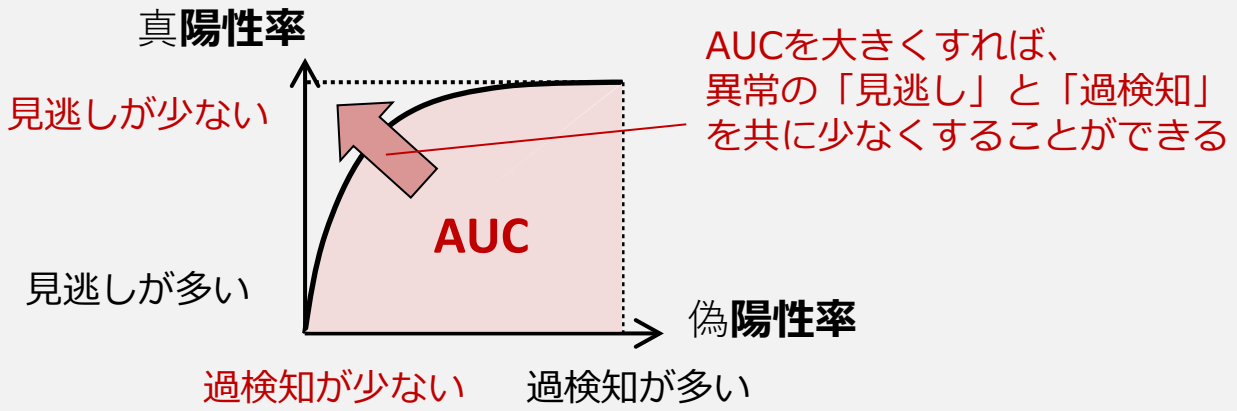
従来の「誤り最小化」に基づく学習では、識別モデルの学習が困難

（少ない異常データを無視して、全データを正常と判定すれば高精度を達成できるため）

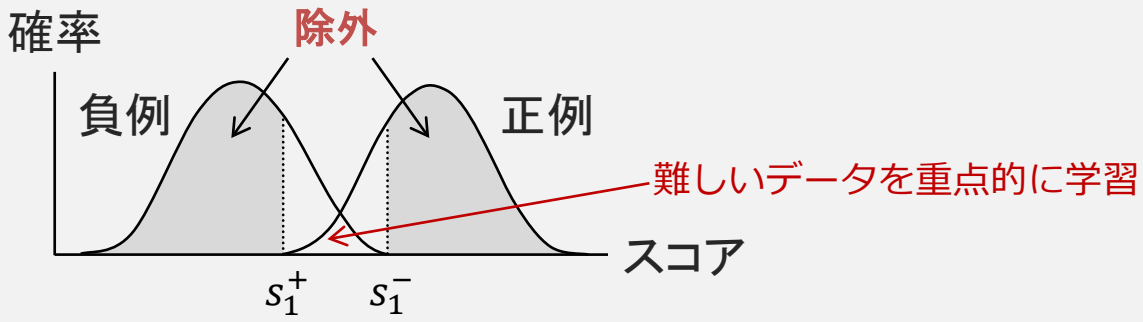
テーマ1：適応的AUC最大化学習（提案手法）

- **学習サンプル選別に基づく、AUC最大化学習**
  - ① AUC最大化により、少ない異常クラスを無視せずに学習
  - ② 識別が難しいサンプルのみを選別することで、効率よく学習

① AUC最大化



② 識別困難なサンプルで学習

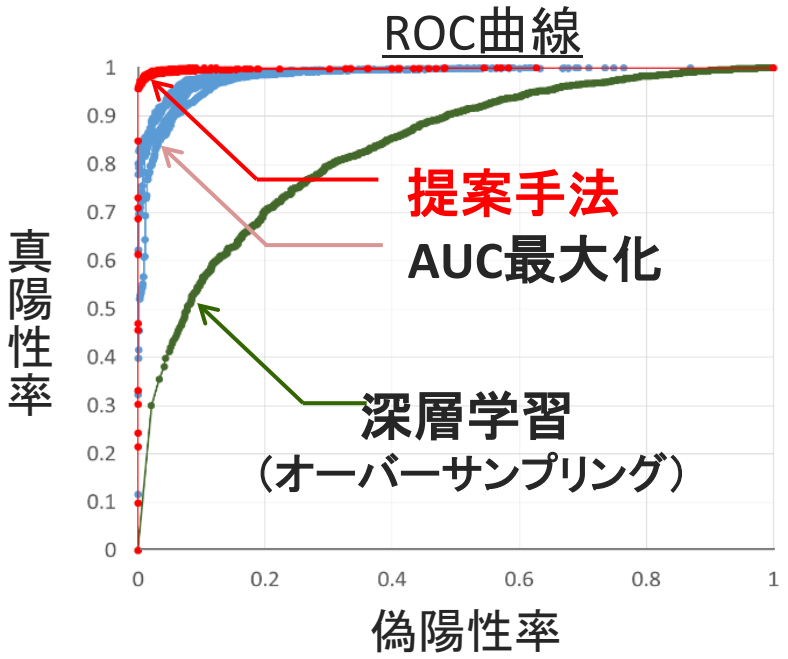


テーマ1：適応的AUC最大化学習（提案手法）

- **不均衡なデータに対して、提案手法の有効性を実証**

CIFAR-10（物体認識）
1クラス500枚を異常とし、残りの
全クラス4500枚を正常として学習

	見逃し率	過検知率
通常学習	0.5%	93%
AUC最大化	0.5%	27%
提案手法	**0.5%**	**3%**



- **工業部品の異常検知など、実データでも有効性を確認**

外観検査向けのNECの機械学習エンジン「RAPID」に搭載

例）屋外インフラ検査

例）精密素材の検品

※画像はgettyimagesの素材画像であり、あくまでイメージです。

テーマ1：弱ラベル学習の高精度化

背景：真のラベルよりも情報量が少ないが、低コストで入手可能な弱ラベルから学習できれば、学習データ作成のコストを大幅削減可能

一般的に知られている弱ラベル学習の例

曖昧ラベル学習

曖昧なラベルで学習

PU学習

一部の正例のラベルのみで学習

補ラベル学習

「～ではない」というラベルで学習

⇒ 物体検知への新クラス追加、部品の検査検品などに応用可能

弱ラベル学習の課題

- 弱ラベルを用いて、真のラベルに対する損失を再構成すると、**訓練損失が下に非有界であるために過学習が生じ、精度が大幅に低下**

テーマ1：弱ラベル学習の高精度化（理論的な成果）

アプローチ

- 1)弱ラベル学習を成功させる「良い」損失関数の条件を設定し、
  2)それを満たすような損失関数を設計することで、高精度な学習を実現

「良さ」の基準

① **Properであること**（クラスの事後確率を正しく推定できること）
- 弱ラベル学習を実現するための十分条件

② **下に有界であること**
- 深層学習のような複雑なモデルで過学習を軽減するために必要

上記①②を満たす損失関数の十分条件を理論導出

$F(\vec{v})$ を凸関数としたとき、$F(\vec{v}) - \max_y (R\vec{v})_y$ が下に有界ならば、
次の形をした損失関数は良い性質①②を満たす：

$$l(\vec{p}, y) = -(R^T \nabla F^*(\vec{p}))_y + F(\nabla F^*(\vec{p}))$$

ただし、$\vec{p}$ はクラス事後確率分布のベクトル表示、$F^*(\vec{p})$ は $F(\vec{v})$ の共役関数。

テーマ1：弱ラベル学習の高精度化（実験結果）

RIKEN AIP Center for Advanced Intelligence Project

導出した条件を満たすように損失関数を補正する正則化手法 Generalized Logit Squeezing (gLS)を考案し、実験的に有効性を確認

- gLSは弱ラベル学習だけではなく、教師あり学習にも適用可能

弱ラベル学習

	CIFAR-10
従来手法① (BC)	29.6%
従来手法② (BC+GA)	36.9%
提案手法　(BC+gLS)	**50.5%**

BC: Backward Correction
GA: Gradient Ascent

⇒ **大幅に精度向上**

教師あり学習

	CIFAR-100	CUB-200	商品認識(実データ)
正則化なし	72.3%	46.6%	84.1%
重み減衰	78.8%	63.2%	85.2%
提案手法 (gLS)	**79.3%** (※)	**64.2%** (※)	**91.2%**

(※) 重み減衰＋gLS

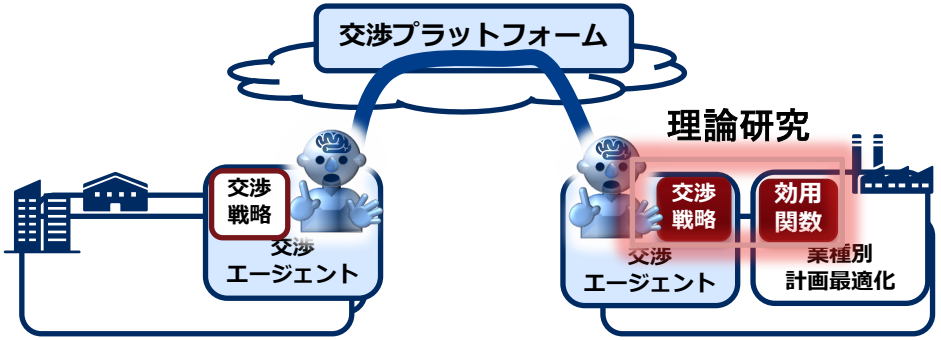
⇒ **教師あり学習でも精度向上**

弱ラベル学習から一般の教師あり学習まで幅広い応用場面で活用

テーマ３：複数AI間の調整にかかわる学習技術の理論解析

- AIが普及した社会において、個別企業内の最適化だけでなく、**AI間での交渉・連携を通して、社会全体を最適化**
- **事業上利用可能**であり、**汎用的に展開可能**な技術の理論研究を進める

- 活動振り返り・成果概要



FY2018-19

様々な事業に適用可能な交渉戦略のための、相互情報量に基づく未知制約付きベイズ最適化

具体的な事業で利用するために、ターゲットを限定

FY2020-21

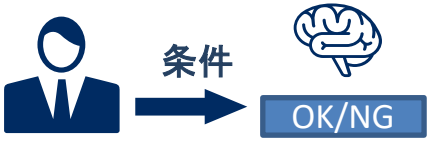
製造・物流で広く使われるスケジュール調整をターゲット設定
交渉の良さを測る効用関数のためのオンライン意思決定

テーマ3：製造・物流で典型的な商取引パターン

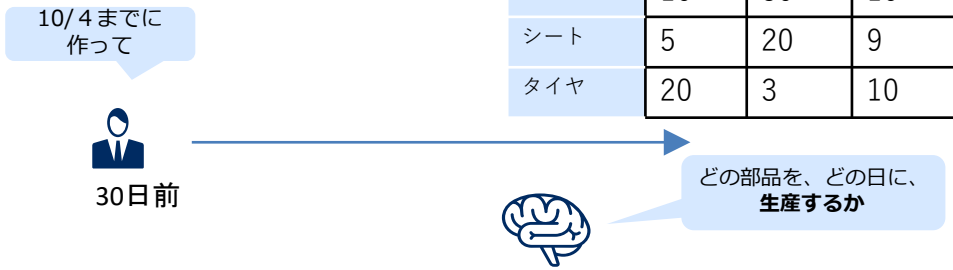
2020年度課題　**製造に多い**

オンライン複数ナップザック問題

・到来した商談の良さを判定するために、自社の資源、使用日を加味し、商談の受理／拒否を判断する。ユーザとのシームレスな対話のために、高速な効用値計算も必要。



課題例

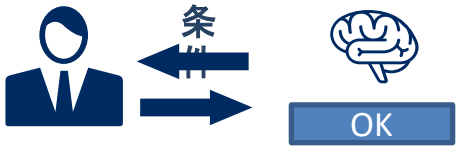


	10/1	10/2	10/3
ハンドル	10	30	10
シート	5	20	9
タイヤ	20	3	10

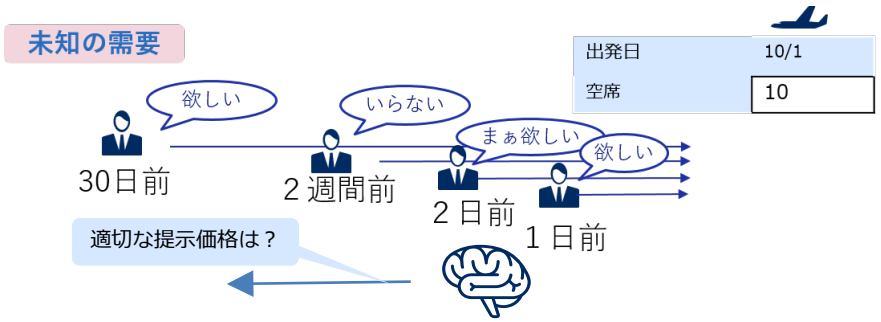
2021年度課題　**物流に多い**

オンラインレベニュー最大化問題

・需要を**未知**とし、在庫、終了時間（消費期限）が決まっている商品に対して逐次的に価格提示を行って、需要を予測し、利益を最大化する



課題例



テーマ3：製造をターゲットにした最適化

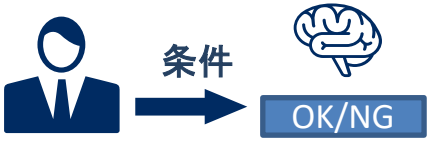
- 製造に多い、不定期に複数到来する（〆切も異なる）顧客要望に対応すべく、オンライン複数ナップザック問題として定式化・解法を考案

2020年度課題

製造に多い

オンライン複数ナップザック問題

・到来した商談の良さを判定するために、自社の資源、使用日を加味し、商談の受理／拒否を判断する。ユーザとのシームレスな対話のために、高速な効用値計算も必要。



課題例

10/４までに作って

30日前

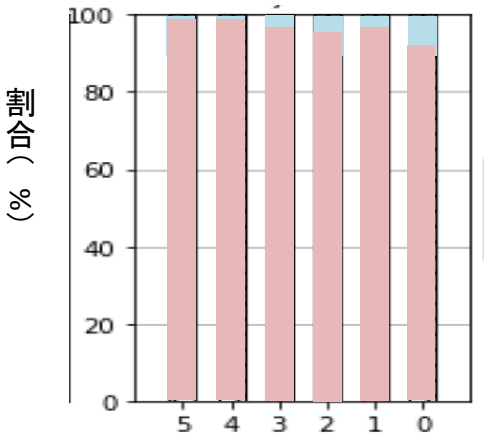
	10/1	10/2	10/3
ハンドル	10	30	10
シート	5	20	9
タイヤ	20	3	10

どの部品を、どの日に、**生産するか**

- 2020年度成果

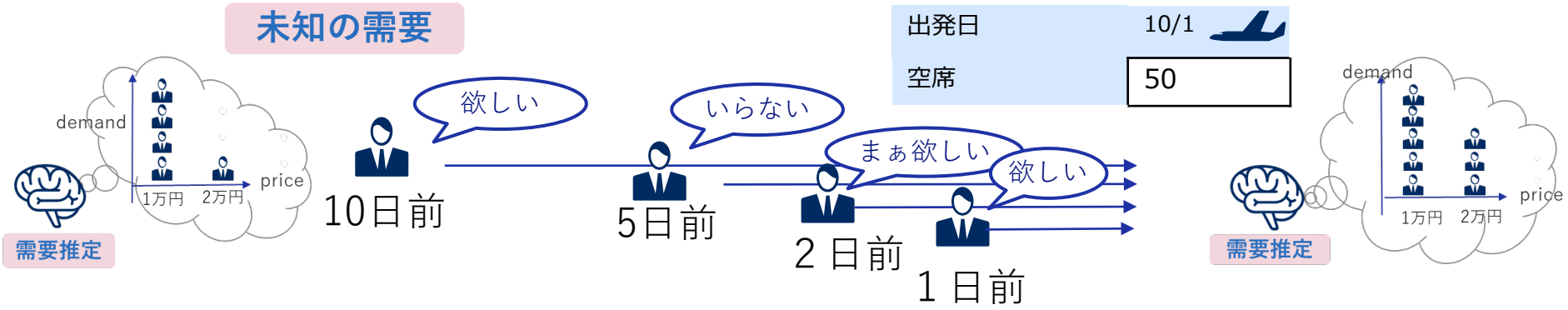
  - 効用関数を確率的ナップザック問題の拡張として定式化
    - 到来する要望・種類が確率的・対応できる容量が可変な場合に対応
  - 最適化手法の算出法、高速な近似解導出手法を考案
  - ルールベース手法（空いている便に順に詰め込む）と比較し開発手法の優位性を確認

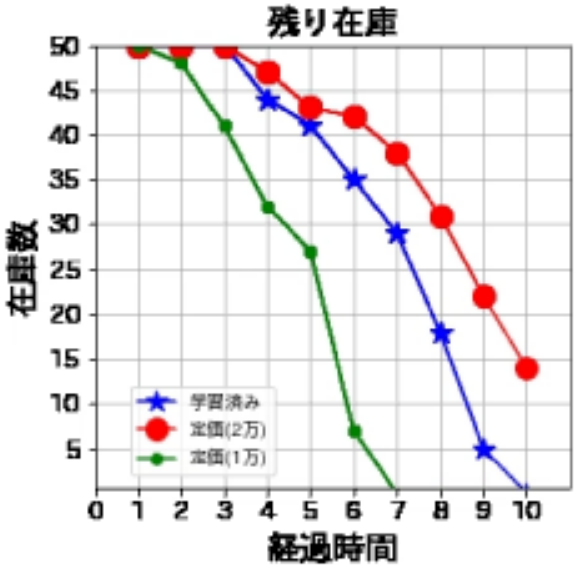
ルールベース（青）／開発手法（赤）
効用関数値が大きかったものの割合

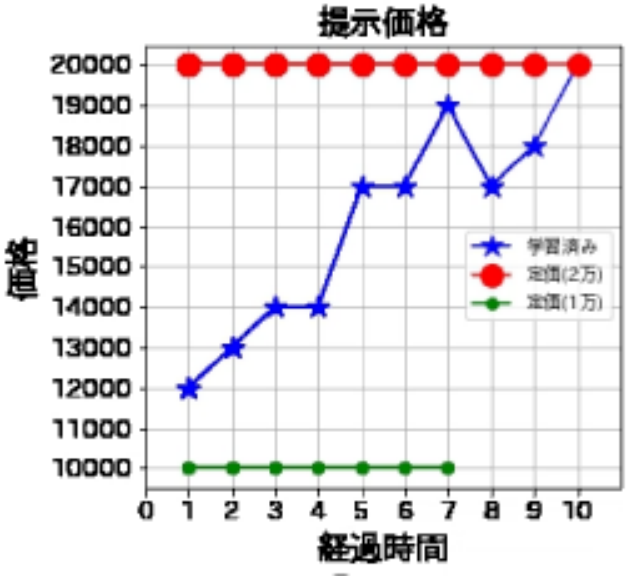


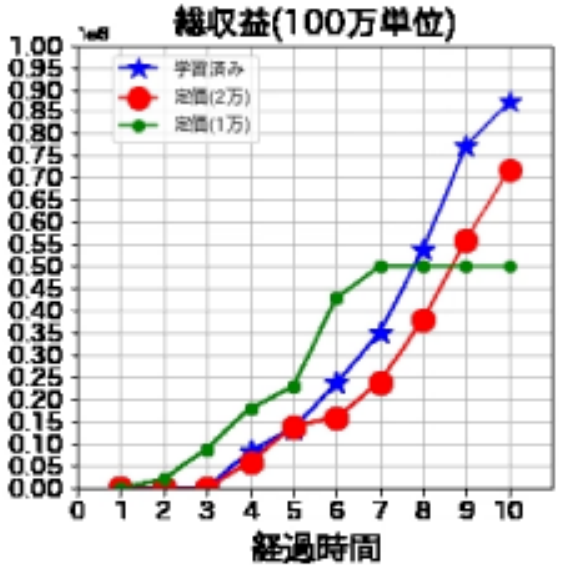
テーマ3：物流をターゲットにした最適化

- 物流に多い需要が未知の場合に対応すべく、動的需要下でのオンラインレベニュー最大化問題として定式化・解法を考案









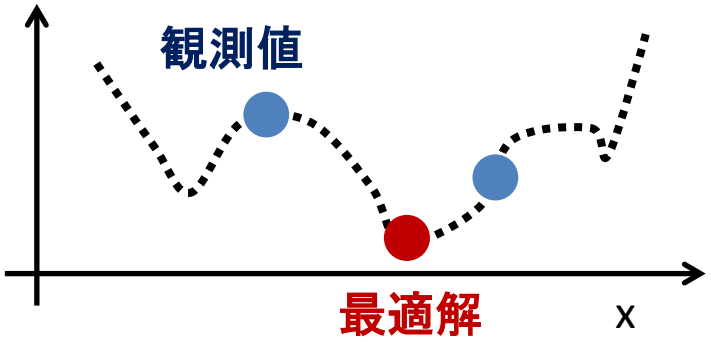
テーマ5：汎用的、パラメータフリーなブラックボックス最適化技術の開発

- **様々な実問題で必要とされるブラックボックス最適化の高度化**
  - **ブラックボックス最適化：具体形が未知（ブラックボックス）な目的関数に対する最適化問題**

関数f(x)の関数形は未知

xの値を決めればf(x)は観測可能

できるだけ少ない観測回数で f(x)を最小化する最適解xに近い解を見出したい

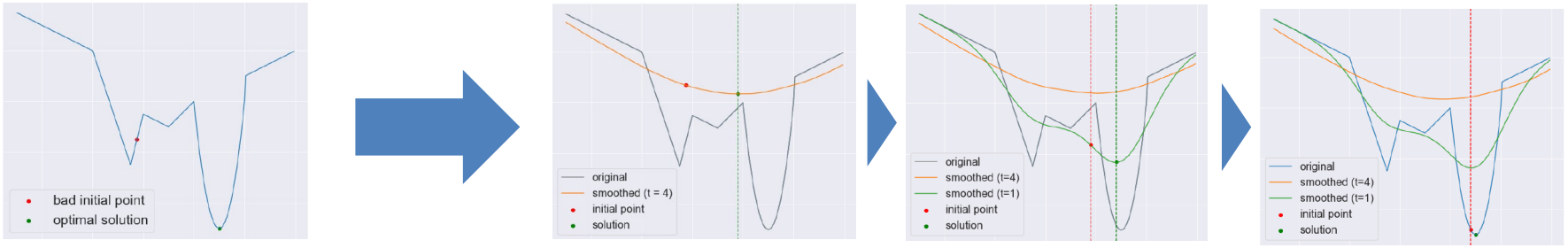


  - **機械学習におけるハイパーパラメータ最適化・モデル選択、実験計画、分子構造解析、創薬、推薦システム、マーケティング自動化など多くのアプリケーションが存在**

- **2021年度テーマを立ち上げ、活動を開始**
- **Zeroth-order optimizationと呼ばれるブラックボックス最適化を解く１つの枠組みを拡張した手法：single loop Gaussian homotopy (SLGH) 法を考案**
  - 様々な問題クラスに対し、理論的に収束性を証明
  - 「ブラックボックス分類モデルに対する敵対的攻撃」タスクで有効性を確認

テーマ５：活動成果概要

- 提案手法の考え方

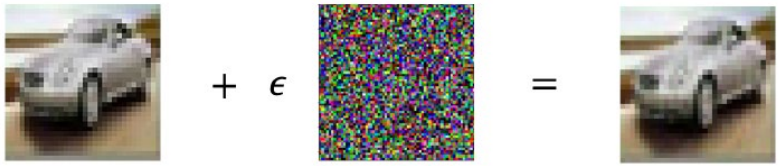


元問題：非凸、非滑らか
そのまま解くのは難しい

スムージングにより
やさしい問題に変換

解の探索と同時にスムージングを弱める
→ 徐々に元問題の解に収束

- 有効性評価
「ブラックボックス分類モデルに対する敵対的攻撃」タスクで有効性を確認



「自動車」
と判定

↑このノイズ
を発見したい

「飛行機」
と誤判定

Table 2. Performance of a per-image attack over 100 images of CIFAR-10 under $T = 10000$ iterations. "Succ. rate" indicates the ratio of success attack, "Avg. iters to 1st succ." is the average number of iterations to reach the first successful attack , "Avg. L_2 (succ.)" is the average of L_2 distortion taken among successful attacks, and "Avg. total loss" is the average of total loss $f(x)$ over 100 samples.

	Methods	Succ. rate	Avg. iters to 1st succ.	Avg. L_2 (succ.)	Avg. total loss
SGD algo.	ZOSGD	88%	**835**	0.076	27.70
	ZOAdaMM	85%	3335	**0.050**	19.15
GH algo.	ZOGradOpt	65%	6789	0.249	39.05
提案手法	ZOSLGH$_r$ ($\gamma = 0.999$)	**93%**	4971	0.248	**14.28**
	ZOSLGH$_d$ ($\gamma = 0.999$)	**93%**	4365	0.188	**14.23**