

(9月9日版)

## 人工知能の概説書

谷口忠大『イラストで学ぶ人工知能概論（改訂第2版）』講談社、2020年（Russell and Norvigのダイジェスト版のような構成で、日本語で読める人工知能の教科書としては、もっとも新しく包括的です。）

メラニー・ミッチェル『教養としてのAI講義ービジネスパーソンも知っておくべき「人工知能」の基礎知識』日経BP、2021年（深層学習などの基本原理や限界を一般向けにわかりやすく紹介している本です。元になる研究論文の書誌情報も充実しています。）

Russell S., and Norvig, P. (2020) *Artificial Intelligence: A Modern Approach (4th edition)*. Pearson. (もっとも有名な人工知能のテキストです。1000ページ以上ある分厚いものです。さまざまな手法がカバーされていますが、この本だけで理解するのは難しいかもしれません。)

## 講義1

Copeland, J. (1993) *Artificial Intelligence: A Philosophical Introduction*. Oxford: Blackwell. (第2次ブーム期までの人工知能の哲学の代表的な概説書。)

Haugeland, J. (1985) *Artificial Intelligence: The Very Idea*. Cambridge, MA.:MIT Press. (こちらも第2次ブーム期までの人工知能の哲学の代表的な概説書。)

ヒューバート・ドレイファス『コンピュータには何ができないかー哲学的人工知能批判』産業図書、1992年（ドレイファスはAI研究批判でもっとも有名な哲学者。この翻訳は1979年出版の原書第2版の翻訳。原書は1992年出版の第3版あり。）

ダグラス・ホフスタッター、ダニエル・デネット『マインズ・アイーコンピュータ時代の「心」と「私」（上）、（下）』TBSブリタニカ、1992年

松尾豊『人工知能は人間を超えるかーディープラーニングの先にあるもの』角川書店、2015年

松尾豊『超AI入門ーディープラーニングはどこまで進化するのか』NHK出版、2019年

## 講義2

### 講義3

French, P. (1990) Subcognition and the Limits of the Turing Test. *Mind*. 99(393): 53-65.

Dreyfus, H., and Dreyfus, S. (1988) Making a Mind vs. Modeling the Brain: Artificial Intelligence Back at a Branchpoint. *Daedalus*. 117: 15-44. (門脇俊介・信原幸弘編『ハイデガーと認知科学』に邦訳あり。)

Harnad, S. (1990) The Symbol Grounding Problem. *Physica D: Nonlinear Phenomena*. 42(1-3): 335-346.

Harnad, S. (1991) Other Bodies, Other Minds: A Machine Incarnation of an Old Philosophical Problem. *Minds and Machines*. 1: 43-54.

Lesvesque, H. (2009) Is It Enough to Get the Behavior Right? *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence*. 1439-1444.

テリー・ウィノグラード、フェルナンド・フローレス『コンピュータと認知を理解する—人工知能の限界と新しい設計理念』産業図書、1989年

サール『心・脳・科学』岩波書店、2015年

ダニエル・デネット「コグニティブ・ホイール—人工知能におけるフレーム問題」『現代思想』1987年4月号、pp. 128-150.

松原仁「一般化フレーム問題の提唱」J・マッカーシー、P・J・ヘイズ、松原仁『人工知能になぜ哲学が必要か—フレーム問題の発端と展開』哲学書房、1990年

### 講義4

アンディ・クラーク『認知の微視的構造—哲学、認知科学、PDPモデル』産業図書、1997年（認知科学における計算主義とコネクショニズムの論争がわかりやすく整理されています。）

ポール・チャーチランド『認知哲学—脳科学から心の哲学へ』産業図書、1997年（1980年代のニューラルネット研究や、その哲学への含意がわかりやすく論じられています。）

### 講義5

柴原一友ほか『機械学習教本』森北出版、2019年（機械学習のさまざまな手法があまり数式を用いずに説明されています。）

大関真之『機械学習入門ーボルツマン機械学習から深層学習まで』オーム社、2016年（数学的な細部は省略されていますが、深層学習の基本的な仕組みがわかりやすく説明されています。）

## 講義6

瀧雅人『これならわかる深層学習入門』講談社、2017年

岡谷貴之『深層学習（改訂第2版）』講談社、2022年

（日本語で書かれた代表的な入門書です。いずれも理系向けの入門書なので、数学の知識がある程度必要となります。）

岡野原大介『ディープラーニングを支える技術ー「正解」を導くメカニズム』技術評論社、2022年

岡野原大介『ディープラーニングを支える技術2ーニューラルネットワーク最大の謎』技術評論社、2022年

（深層学習に関連するさまざまな手法が紹介されています。上記2冊よりも図による直観的な説明が豊富です。）

## 講義7

深層学習による画像認識に関しては、上の瀧本や岡谷本で解説されています。

ジェリー・フォード『精神のモジュール形式ー人工知能と心の哲学』産業図書、1985年

（講義で言及した心のモジュール性に関するフォードの著作です。）

## 講義8

柴原一友、藤本浩司『続 AIにできること、できないことーすっきり分かる「最強AI」のしくみ』日本評論社、2019年

（画像認識のResNet、自然言語処理のBERT、ゲームのAlphaZeroの基本的な発想がわかりやすく解説されています。）

中山光樹『機械学習・深層学習による自然言語処理入門ーscikit-learnとTensorFlowを使った実践プログラミング』マイナビ、2020年

## 講義9

Hayes, P. and Ford, K. (1995) Turing Test Considered Harmful. *Proceedings of the 14th International Joint Conference on Artificial Intelligence*. 1: 972-977.

ロドニー・ブルックス『ブルックスの智能ロボット論—なぜMITのロボットは前進し続けるのか?』オーム社、2006年

## 講義10

大坪直樹ほか『XAI（説明可能なAI）—そのとき人工知能はどう考えたのか?』リックテレコム、2021年

デイヴィッド・サンプター『数学者が検証！アルゴリズムはどれほど人を支配しているのか?—あなたを分析し、操作するブラックボックスの真実』光文社、2019年

## 補足資料1

L・ゴールドシュレーガー、A・リスター『計算機科学入門（第2版）』近代科学社、2000年

Noam Nisan, Shimon Schocken『コンピュータシステムの理論と実装—モダンなコンピュータの作り方』オライリー・ジャパン、2015年（原書は2021年出版の第2版あり）

チャールズ・ペゾルド『CODE—コードから見たコンピュータのからくり』日経BPソフトプレス、2003年（原書は2022年出版の第2版あり）

## 補足資料2

## 補足資料3

マルコフ決定問題、強化学習、深層強化学習については、上記岡野原本の2巻で紹介されています。ただし、これだけだと十分に理解するのは難しいかもしれません。

Morales, M. (2020) *Grokking Reinforcement Learning*. Manning.

（英語ですが、強化学習の解説書にいろいろ目を通した中では、これが一番わかりやすかったです。）

## 補足資料4