

実践知能アプリケーション構築プラットフォームPRINTEPS

PRINTEPS: A Platform to Develop Practical Intelligent Applications

山口 高平†

Takahira Yamaguchi †

†慶應義塾大学 理工学部

† Faculty of Science and Technology, Keio University

要旨

本稿では、人と機械がマルチモーダル（知識、対話、表情、視線、姿勢、動作、環境）でインタラクションをとることにより、人の知能と機械知能が互いに進化し続けていく「知能共進化」の枠組みについて述べ、多重知識ベース、機械学習、音声対話、人センシング、環境センシング、マニピュレーションという 6 種類の知能を実現するためのソフトウェアモジュールライブラリをもつプラットフォーム PRINTEPS(PRACTICAL INTELLIGENT aPPlicationS)の構成法を提案し、ケーススタディとして、PRINTEPS をマルチロボット連携による模擬喫茶店に適用した結果について述べる。

1. はじめに

近年、Nao や Pepper のようなコミュニケーションロボットが我々の身近になってきて、人工知能 (AI) が人の職業を奪うような議論も出てきている。しかしながら、例えば、Pepper の対話アプリを稼働させると、Pepper は対話発言権をユーザになかなか譲らず、自分の発言に人を強引に引き込もうとする。これは、Pepper は、想定シナリオ範囲内でのみ対応可能であり、ユーザに発言を許してしまうと、コミュニケーションが途端に破綻するためである。人と自由なコミュニケーションができる AI を実現するには、課題は山積しており、コミュニケーション型 AI の方が、ゲームやクイズなど特定のタスクで人に迫る自律型 AI より、実現は困難である。さらに、画像センシング（画像認識、音声認識）の研究分野では、Deep Learning（深層学習）の精度が、従来の機械学習（SVM、CRF、ベイジアンネットなど）の精度を凌駕し大きな注目を集めているが、Deep Learning を意味処理に適用しても、現状では大きな効果はなく、深い意味処理の実現には課題が山積している。CYC という世界最大規模のオントロジー（概念間の上位下位関係と意味関係）は、30 年以上手作りりで開発が続いており、深い意味処理の自動化は多くの困難が伴う。ただ、手作りりで開発されてきた良質のオントロジーが整備され、意味処理の利用方法を検討する時代が到来しているともいえ、上述のコミュニケーション型 AI への適用はチャレンジといえる。

以上の事から、近未来の AI は、知覚処理と意味処理を統合するコミュニケーション型 AI、さらにそれを発展させて、人と AI が相互に連携して、お互いの知能（＝知識・データ＋知的処理）を高めていく

「知能共進化 (Co-evolution of Intelligence)」型 AI の研究を開始した。以下においては、知能共進化の提案背景、そのプラットフォーム PRINTEPS (Practical INTElligent aPPlicationS) の設計、マルチロボット連携喫茶店の実現に PRINTEPS を適用した結果と評価について述べる。

2. PRINTEPS の設計方針

図 1 に、人の知能、機械（ロボット、キャラクターエージェントなど）の知能、PRINTEPS が提供する知能モジュール群の対比を示す。

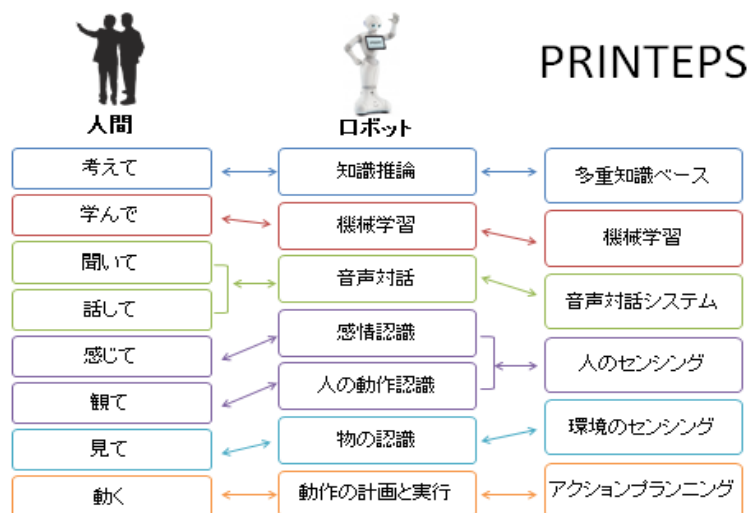


図 1 人の知能と機械知能と PRINTEPS の対比

PRINTEPS は、人と協働する知能アプリケーションを開発するためのプラットフォームであり、多重知識ベース（知識を利用した推論）、機械学習（スキルや知識の向上）、音声対話システム（音声対話）、人のセンシング（人の感情や動作認識）、環境センシング（物体の認識）、アクションプランニング（ロボットの手足の動作の計画と実行）の 6 種類の知能モジュール群をもつ。さらに、これらの知能モジュール群を統合し、知識・対話・表情・視線・姿勢・動作・環境の 7 モードから成るマルチモーダルインタラクションを実装し、人も機械も互いに学びあって共に進化していくシステムを実現するものである。

多重知識ベースは、常識オントロジー、専門分野の領域オントロジー、タスク処理手順であるワークフロー、状況判断を言及するルールベース、判断根拠を提示するゴール木から構成され、日本語 Wikipedia、言語資源、既存オントロジー、関連 Web ページなどから、初期の多重知識ベースを開発する。その後、マルチモーダルインタラクションにおけるユーザの指摘、および、知識ベース全体の首尾一貫性テストによって、多重知識ベースを進化させていく。

機械学習モジュールは、データの前処理、機械学習、学習結果の後処理に関するソフトウェアモジュール群を整備し、観察データから規則性を一括学習するか、事前に与えた業務ルールを洗練させる。例えば、ロボット喫茶店において「IF 来店客が大きな荷物を持つ THEN 大きな荷物を空きスペースに置かせる」という来店客対応ルールがあり、そのルールを来店客が受容するケースとしないケースが分かれば、カメラやセンサーを通して得られた情報をルール条件部に追加し、そのルールを分化させる。

音声対話システムでは、多重知識ベースに基づき、ユーザの理解度を推定しながら円滑な対話を推進するための対話知識ベースの開発、人・環境センシングの情報から、視線・表情・姿勢・動作特徴・対人距離などからユーザの会話参加態度を推定するモジュールの開発、話す内容の順序、多人数において話しかける人の選定、雑談の展開の仕方などを管理する、マルチモーダル対話モデルを開発する。

人センシングでは、人の表情解析、人の視線推定、人の検出と追跡、人の姿勢推定、人の行動を認識と予測などのモジュール群を開発する。

環境センシングでは、単眼・多眼カメラ装置から、色・距離・温度を推定するとともに、環境・物体・人物の様々な形状と特徴量を抽出し、人センシングと併せて、非言語情報インタラクション基盤とする。

アクションプランニングには、物の把持計画、障害物を避けながら移動する計画など、マニピュレーションと移動に関するモジュール群、および、手差しなどの身振り、感情に関わる手振りなどのモジュール群を開発する。

図 2 に、エンドユーザが記述する「サービス」「(ビジネス) プロセス」、開発者が記述する「(ソフトウェア) モジュール」「多重知識」「データ」という 5 階層から成る、実践知能アプリケーション開発プラットフォーム PRINTEPS のシステム構成を示す。左半分が、教諭とロボットの連携授業、右半分が、ロボット喫茶店を想定したシステム構成である。

知能実践アプリケーション開発プラットフォーム PRINTEPS



5

図 2 PRINTEPS のシステム構成

3. ロボット喫茶店

以下、ロボット喫茶店の実現方法として、ロボットによる入店時挨拶プロセス（図3）を例にとり説明する。図3において矩形は、入店時挨拶プロセスを構成するモジュールを、矢印はモジュールのフローを、矢印のラベルはモジュールの入出力をそれぞれ示している。入店時挨拶プロセスでは、最初に発話ロボット検索を行い、検索された発話ロボットが年齢認識モジュールにより顧客の年齢を推定する。次に、年齢による人のタイプ判定ルールを参照し、人モデルを構築する。同時に、発話ロボット検索結果よりロボットモデルを構築し、人モデルとロボットモデルより、喫茶店モデルを構築する。その後、喫茶店モデルと人のタイプから挨拶内容を決定するルールを参照し、入店時挨拶内容取得モジュールにより、発話テキストを得る。最後に、発話モジュールにより、発話テキストを音声合成により発話するといった流れとなっている。

ここで、各モジュールの入出力には、図4に示す喫茶店業務クラス階層を用いている。入店時挨拶プロセスで必要となるクラスとしては、人(Person)クラスとそのサブクラスとして顧客のタイプを判別するために用いる大人(Adult)や子供(Child)などのクラス、ロボット(Robot)クラス、喫茶店(Cafe)クラスを定義した。人クラスのデータタイププロパティとしては、人の年齢を表すageプロパティ(定義域: Person, 値域: int)を定義した。ロボットクラスのデータタイププロパティとしては、発話内容を表すためのgreetingプロパティ(定義域: Robot, 値域: String)を定義した。喫茶店クラスのオブジェクトプロパティとしては、喫茶店にいる顧客との関係を表すためのhasClientプロパティ(定義域: Cafe, 値域: Person)、喫茶店にいるロボットとの関係を表すためのhasRobotプロパティ(定義域: Cafe, 値域: Robot)などを定義した。

図5に、年齢による人のタイプ判定ルールと人のタイプから挨拶内容を決定する来店客対応ルールを示す。

年齢による人のタイプ判定ルールとしては、12歳以下の子供(YoungChild)クラスをタイプとして持つ、13から20歳までの人は少年(TeenAge)クラスをタイプとして持つといったようなルールを定義している。人のタイプから挨拶内容を決定するルールとしては、例えば、大人と少年に対しては「いらっしゃいませ」、小さな子どもに対しては「お母さんは、どこにいるの?」といったような挨拶内容を決定するルールを定義している。

ロボット喫茶店は、来店時挨拶→席案内→注文→飲み物作成→配膳→会計→見送りとなり、現在、連続的ではなく断片的にしか実行できない状況であるが、その様子を図6に示す。

入店時挨拶プロセス

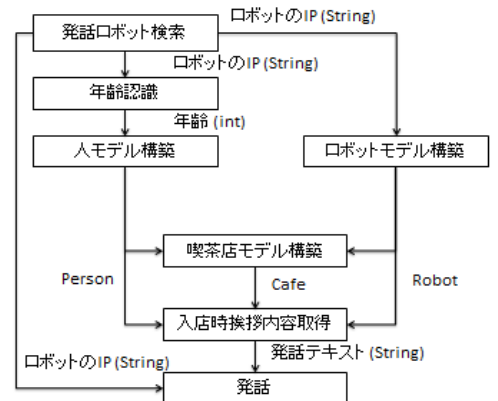


図3 入店時挨拶プロセス

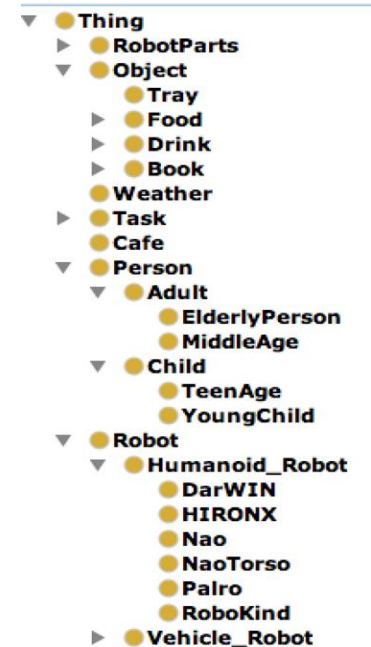


図4 喫茶店業務クラス階層

- 年齢による人のタイプ判定ルール
- Person(?a), age(?a, ?i), lessThan(?i, 13) -> YoungChild(?a)
 - Person(?a), age(?a, ?i), greaterThan(?i, 12), lessThan(?i, 20) -> TeenAge(?a)
 - Person(?a), age(?a, ?i), greaterThan(?i, 19), lessThan(?i, 70) -> MiddleAge(?a)
 - Person(?a), age(?a, ?i), greaterThan(?i, 69) -> ElderlyPerson(?a)
- 人のタイプから挨拶内容を決定するルール
- Cafe(?x), Robot(?z), YoungChild(?y), hasClient(?x, ?y), hasRobot(?x, ?z) -> greeting(?z, "お母さんは、どこにいるの?")
 - Cafe(?x), Robot(?z), TeenAge(?y), hasClient(?x, ?y), hasRobot(?x, ?z) -> greeting(?z, "いらっしゃいませ")
 - Adult(?y), Cafe(?x), Robot(?z), hasClient(?x, ?y), hasRobot(?x, ?z) -> greeting(?z, "いらっしゃいませ")

図5 来店客対応ルール



図6 ロボット喫茶店の一連の流れ

4. おわりに

本稿では、知能アプリケーションをアジャイルに開発するプラットフォーム PRINTEPS (PRacticalINTElligent aPplicationS)の構成法とロボット喫茶店への適用結果について述べた。文献[2]では、小学校教諭とロボットが連携して授業を実践し、子供達の興味関心が向上した結果が得られた。しかしながら、このようなロボットアプリの開発は、実機動作の確認が必要であるため開発コストは大きい。今後、PRINTEPS の SW モジュールの整備を進め、アジャイルに知能ロボットアプリを開発できる環境に進展させていきたいと考えている。

参考文献

- [1]森田武史, 山口高平: PRINTEPS アーキテクチャの構成と実践, 第 29 回人工知能学会全国大会,114-1 (2015)
- [2]森雄一郎, 菅陽哉, 菅原優, 山口高平: 知能共進化アーキテクチャに基づく小学生教育支援ロボット, 情報処理学会第 77 回全国大会, 1P-09 (2015)

謝辞

本研究は、科学技術振興機構(JST) 戦略的想像研究推進事業(CREST)「実践知能アプリケーション構築フレームワーク PRINTEPS の開発と社会実践」の支援によって実施した。