

オープンチャットとロボットの連係による ティーチングアシスタントとの協働システム

A Collaborative System with Teaching Assistants
based on Linkage of Open Chat and Robots

渥美雅保 村田祐樹 安川葵
Masayasu Atsumi Yuuki Murata Aoi Yasukawa

創価大学工学部情報システム工学科

Dept. of Information Systems Sci., Faculty of Sci. and Eng., Soka University

This paper describes a collaborative system with teaching assistants based on linkage of open chat and robots. The system consists of a chat bot, chat clients on web browsers for students and teaching assistants and two mobile robots with their remote brain servers which are linked with a chat server. The characteristic functions of a chat bot are selective intervention based on a student utterance classifier and a response reasoning based on an attention-based LSTM encoder-decoder neural network named CoMNet. The characteristic function of mobile robots is interactive learning of person re-identification for collaborative teaching support. These functions are evaluated using datasets collected during the system operation in a software practice class.

1. はじめに

チャットボットやコミュニケーションロボットの教育の分野における利用が注目されるようになってきている [Maderer 16, Mubin 13]. 我々は、オープンチャットとロボットの連係による演習授業支援のためのティーチングアシスタント (TA) との協働システム SOTARO (SOciable Teaching Assistance RObot) を開発して運用をはじめている [渥美 17, 村田 17a, 村田 17b, 安川 17]. 本システムは、受講学生と TA、及びチャットボットによる三者のオープンチャットとそれと連係して教室内を巡回して学生とインタラクションをする複数台の移動ロボットから構成され、TA と役割分担して学生の学習を支援するとともに機械学習を通じて機能を獲得する仕組みを備えている。本システムの主な特徴的機能は次の 3 つである:

(1) オープンチャットへのボットの選択的介入による自動応答: 学生と TA のチャットに選択的にボットが参加して学生の発言に対して応答を自動的に返す。どの発言に自動応答するかの判別のために、ランダムフォレストによる発言分類に基づく方法を開発して用いている。また、自動応答推論に関して、CoMNet (Conversation Memory Network) と名付けた会話コンテキストと文内フレーズへの注意を組み込んだ LSTM (Long Short-Term Memory) エンコーダ・デコーダ型ニューラルネットワークを開発して用いている。チャットはテキストで行うが、音声認識・合成の利用も可能である。

(2) 移動ロボットの認識-制御とその学習のリモートブレインでの TA との協調的実行: 移動ロボットは RGB-D 映像と音声、及びテキスト・制御コマンドをサーバのリモートブレインとの間で通信することにより遠隔で制御される。TA が移動ロボットを介して学生とインタラクションをする際に学生を識別して関連情報を参照するために、人物再同定を人間参加 (Human-in-the-loop) 型のインタラクティブな教師付き機械学習で行う方法を開発して用いている。

(3) オープンチャットと移動ロボットの対面コミュニケーションの連係: チャットの内容をロボットで選択的に音声合成して発信、また、ロボットと学生との音声による対面コミュニケーションを選択的にチャットや別のロボットで発信することにより、実空間とサイバー空間でのコミュニケーションを連係する。

連絡先: 渥美雅保, matsumi@t.soka.ac.jp

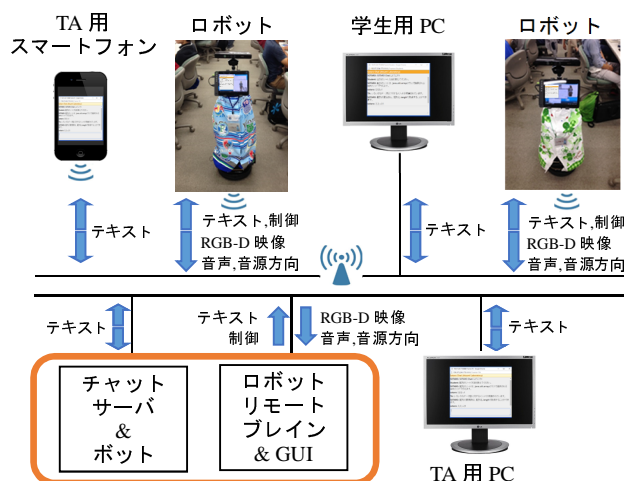


図 1: システムの構成

本論では、本システムの全体構成、特徴的機能である発言分類、応答推論、人物再同定、TA との協働のシナリオ、及び大学のソフトウェア演習の授業での運用で得られたデータを用いた性能評価実験について述べる。このうち、発言分類と人物再同定については [村田 17a, 村田 17b, 安川 17] に詳しい記述があるので、主に CoMNet による応答推論について説明する。

2. システムの構成

2.1 全体構成

図 1 にシステムの構成を示す。本システムは、チャットサーバとボット、ウェブブラウザ上のチャットクライアント群、及び 2 台の移動ロボットとそれぞれのリモートブレインサーバから構成される。チャットクライアントは、学生の PC, TA の PC, TA のスマートフォン、及び移動ロボットで動作し、それらクライアント間並びにサーバで動作するボットとの間のオープンチャットはチャットサーバにより管理される。チャットボットの機能は、ランダムフォレストによる学生の発言分類に基づく選択的介入判定と CoMNet による学生の発言に対する応答推論とから構成される。移動ロボットは、TurtleBot をベースに構築され、RGB-D センサとしては Kinect が搭載されてい

る。リモートブレイクサーバは、RGB-D 映像と音声情報をロボットから受信して、TA との協調のもとで制御コマンドと発話テキストをロボットへ送信する。RGB-D 映像からは、インタラクティブな教師付き機械学習による人物再同定と深度情報を用いた空間認識を行う。ただし、移動制御に関しては、安全で正確な移動のため、現状、TA による GUI を用いた遠隔制御に任せている。音声対話については、チャットとは別のシステムが用意され、TA による対話及びルールベースに基づく対話の機構が組み込まれている。チャットサーバとリモートブレイクサーバは連携してコミュニケーションサービスを提供する。

2.2 チャットボットによる自動応答の判定

オープンチャットでの学生の発言は質問や呟きなど様々であり、それらにはボットでの自動応答に適したものと TA が対面に対応すべきものがある。そこで、学生の発言が自動応答に適した「自動応答タイプ」か TA が対面に対応すべき「TA 応答タイプ」かを判定する [安川 17]。そのために、発言をランダムフォレストを用いて「話題カテゴリ」並びに「期待応答タイプ」により分類する。ランダムフォレスト分類器では、発言の素性として、名詞、動詞、副詞、連体詞、助詞、助動詞の Bag of Words, 漢字, ひらがな, カタカナ, 数字, 記号, アルファベットの文字種割合, 及び文末の単語 3-gram を連結したものをを用いる。話題カテゴリとしては、Java のソフトウェア演習の授業支援用のため、「Java 基礎知識」「エラー・トラブル」「授業運営」「雑談一般」の 4 つを設定する。また、期待応答タイプとしては、「原因」「方法」「定義」「事実・意見」「不定」の 5 つを設定する。ここで、「不定」は、呟き等期待する応答がはっきりしない発言のタイプとする。そして、発言の話題カテゴリ、期待応答タイプ、もしくはそれらの組み合わせにより、自動応答タイプか TA 応答タイプかを判定し、自動応答タイプの場合は 3. で述べる CoMNet により応答を推論して自動応答し、TA 応答タイプの場合には TA の対応予告をチャットで流して TA の PC 及びスマートフォンではそれを音声合成して通知する。これにより、例えば、話題カテゴリ「エラー・トラブル」の発言は TA 応答タイプでその他は自動応答タイプ、期待対タイプ「原因」「方法」の発言は TA 応答タイプでその他は自動応答タイプ、また、話題カテゴリと期待応答タイプを組み合わせ「雑談一般」を除く「原因」「方法」の発言は TA 応答タイプ等、発言をボットの自動応答と TA の対面応答に仕分ける。

2.3 人間参加型機械学習に基づく人物再同定

演習教室を巡回する移動ロボットが、コンピュータに向かって演習中の学生を再同定するために、インタラクティブな人間参加型の教師付き学習により人物識別器をオンラインで生成する [村田 17a, 村田 17b]。図 2 に、人間参加型機械学習による人物再同定手法の枠組を示す。本手法では、画像から HOG(Histograms of Oriented Gradients) 特徴を求めて

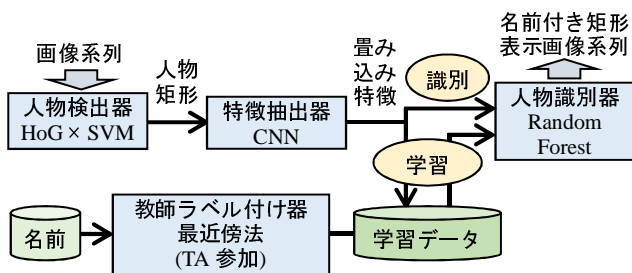


図 2: 人間参加型機械学習による人物再同定

SVM(Support Vector Machine) により人物の上半身矩形を検出し、CNN(Convolutional Neural Network) により人物矩形の畳み込み特徴を計算する。CNN には CaffeNet を用い、全結合層 fc6 の 4096 次元を特徴としている。人物再同定器の学習は、これら人物矩形の畳み込み特徴に対する最近傍法に基づく人間参加型教師ラベル付けとそれら教師付きデータからのランダムフォレスト分類器の生成とからなる。人間参加型教師ラベル付けでは、最近傍法を用いたクラスタリングに TA が参加して GUI を介してクラスタの統合・調節、及びクラスタの名前付けを行う。人物再同定では、人物矩形の畳み込み特徴からランダムフォレスト分類器により学生の名前を求めて名前付きの矩形を画像に表示する。ランダムフォレスト分類器の学習・更新は、授業中に逐次収集する人物画像データを用いて一定時間間隔で実施される。

3. CoMNet による応答の学習と推論

学生の発言に対する応答は CoMNet により推論される。CoMNet は、会話コンテキストと文内フレーズへの注意を組み込んだ LSTM エンコーダ・デコーダ型ニューラルネットワークである。図 3 に CoMNet の構成の概略を示す。CoMNet は、入力発話文エンコーダ LSTM, 応答発話文デコーダ LSTM, 文内フレーズ注意ネットワーク, 及び会話コンテキスト注意ネットワークから構成される。入力発話文エンコーダと応答発話文デコーダの単語列埋込み層には Word2Vec により事前学習した単語ベクトルが組み込まれる。また、応答発話文デコーダの単語列生成層では同じ単語ベクトルが学習時の初期値として用いられる。入力発話文エンコーダには、単語列を前からと後ろからエンコードする双方向 LSTM(Bidirectional LSTM)[Graves 05] か後ろから単方向にエンコードする LSTM[Gers 00] が用いられる。文内フレーズ注意ネットワークは、エンコーダ LSTM の隠れ状態系列 $H_e = [h_{e_1}, \dots, h_{e_L}]$ とデコーダ LSTM の各隠れ状態 $h_{d_j} (j = 1, \dots, M)$ に対して [Bahdanau 15] と同様に 2 層のフィードフォワードネットワークにより注意度 a_{e_i, d_j}^p を計算し、注意度による H_e の重み付け和 $\sum_{i=1}^L a_{e_i, d_j}^p \times h_{e_i}$ を求めてデコーダ LSTM に入力することにより応答発話文デコーダでの単語生成を制御する。ここで、 L は入力発話文の単語列長、 M は生成される応答発話文の単語列長である。

会話コンテキスト注意ネットワークは、学生の発言に対するボットの応答生成に、学生とボットの発言履歴を反映させるネットワークで、学生とボットの発言履歴を管理する最大長 N のキューとそれを用いて注意コンテキストを計算するネットワークから構成される。学生またはボットの発言を $u_t = (s_t, p_t)$ とする。ここで、 $s_t = [w_{t_1}, \dots, w_{t_L}]$ は発話文の単語列、 p_t は発話者で学生かボット、 t は発話時点を与えるインデクスである。

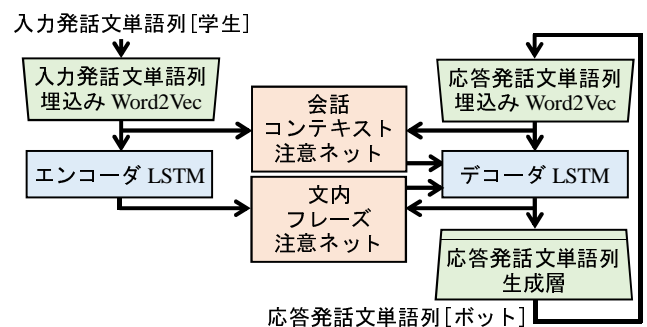


図 3: CoMNet の構成

このとき、発話系列に対して、次のように注意コンテキストを計算する。まず、発話文の単語列を入力または応答発話文単語列埋込み Word2Vec で埋込んだ単語ベクトル列 $\{v_{t_1}, \dots, v_{t_L}\}$ に対してそれらの和 $\hat{v}_t = \sum_{k=t_1}^{t_L} v_{t_k}$ を求める。そして、 \hat{v}_t と p_t のペアにその影響度 $e_t = 1.0$ を付与した $((\hat{v}_t, p_t), e_t)$ を発話履歴キューにプッシュする。また、キュー内発話履歴が最大キュー長 N を超えたならば最も古い発話履歴を 1 つポップする。発話者 p_t がボットの場合は、これらプッシュとポップによる発話履歴キューの管理のみを行う。発話者 p_t が学生の場合は、発話履歴キューから注意コンテキストを計算する。発話履歴キュー内の長さ $n (n \leq N)$ の発話履歴を $[q_t, \dots, q_{t-n+1}]$, $q_k = ((\hat{v}_k, p_k), e_k)$ とする。ここで、 t は現時点のインデクスである。このとき、現時点 t の学生発話の (\hat{v}_t, p_t) と各時点 $t' (t > t' \geq t - n + 1)$ の $(\hat{v}_{t'}, p_{t'})$ との類似度 $s_{t'}$ を式 (1) により計算する。

$$s_{t'} = \max\left(0, \frac{\hat{v}_t \cdot \hat{v}_{t'} + \delta_p(p_t, p_{t'})}{\sqrt{|\hat{v}_t|^2 + 1} \sqrt{|\hat{v}_{t'}|^2 + 1}}\right) \quad (1)$$

ここで、 $\delta_p(p_t, p_{t'})$ は p_t と $p_{t'}$ が同じときに 1、そうでないとき 0 を返す。そして、類似度 s と影響度 e とから式 (2) により関連度 r を求めて、同時に影響度 e を式 (3) により更新する。

$$r_{t'} = \begin{cases} s_{t'} & \text{if } s_{t'} \leq e_{t'} \wedge t > t' \geq t - n + 1 \\ e_{t'} & \text{if } s_{t'} > e_{t'} \wedge t > t' \geq t - n + 1 \end{cases} \quad (2)$$

$$r_t = e_t \quad (3)$$

注意コンテキストは関連度と発話履歴から式 (4) のフィードフォワードネットワークにより計算され、デコーダ LSTM に隠れ状態の初期値として渡される。

$$\tanh\left(W \sum_{k=t}^{t-n+1} \left(\frac{r_k}{\sum_{i=t}^{t-n+1} r_i} * \hat{v}_k\right)\right) \quad (4)$$

ここで、 W は重みパラメータ行列である。

CoMNet の学習は、発話文と応答文のペア集合を用いて end-to-end に実行される。学習時には、応答発話文デコーダでは教師応答文の単語列を入力し、生成される応答発話文の単語分布列との間の誤差を逆伝播する。

4. 協働シナリオと運用

SOTARO の運用を行っているソフトウェア演習の授業は、PC 教室で UML と Java を用いたオブジェクト指向プログラミングを学ぶ授業で、受講生は 40 名程度である。そこでの協働支援シナリオは次のようなものである。移動ロボットは教室に入ってくる学生を人物検出して人物特徴を抽出して、人物識別器の学習を始める。人物識別器の学習は実時間で自動抽出する人物特徴をもとに一定時間間隔で人間参加型の機械学習に基づき行われる。これにより映像内の学生は名前でラベル付けされ、TA はロボットを介するリモートコミュニケーションで名前を利用することが可能となるとともに、学生の学習履歴等を映像に紐づけて利用することが可能となる。学生からの質問等は TA の呼び止めに加えてチャットもしくはロボットへの話しかけを通じてなされる。チャットでは、「自動応答タイプ」と判定された発言にはボットが応答し、「TA 応答タイプ」と判定された発言には TA が PC でチャットもしくは出向いて対面で応答する。また、ボットの自動応答が適切でない場合には TA

がチャットで補足をする。この TA による補足情報は、オフラインでのボットの応答生成学習に利用される。

移動ロボットを用いた運用は 2015、2016 年度の秋セメスターに合わせて 30 回、オープンチャットとロボットの係による運用は 2016 年度の秋セメスターに 15 回実施し、運用中に収集したデータを用いて機能の改良、拡張を続けている。

5. 運用データを用いた実験評価

ソフトウェア演習での運用で用意及び収集したデータを用いて、発言分類、応答推論、人物再同定の実験評価を行った。

5.1 発言分類実験

チャットでの学生の発話文と事前に用意した Java の FAQ 等の質問文からなる 1816 個の文を学生の発言文データセットとして、ランダムフォレストによるそれらの話題カテゴリと期待応答タイプの分類性能を評価した。話題カテゴリ数と期待応答タイプ数は 2.2 で述べたようにそれぞれ 4 つと 5 つである。また、2.2 で述べた発言の素性はこのデータセットに対して 2562 次元である。これらデータのうち、学習には 1452 文を用いて 5 分割交差検証を行い、テストには 364 文を用いた。グリッドサーチにより定めたランダムフォレストの決定木数は、話題カテゴリ分類器では 500、期待応答分類器では 100 である。テストデータに対する正解率は、話題カテゴリ分類で 0.835、期待応答タイプ分類で 0.819 であった。

5.2 応答推論実験

チャットログと Java の FAQ 等からの入力発話文、応答発話文それぞれ 1834 個の計 3668 個の文をデータセットとして、CoMNet による応答推論の性能評価を行った。データセットの語彙数は 6554、文の平均単語数は 23.4 単語である。CoMNet に組み込む単語ベクトルは C-BoW (Continuous Bag of Words) によりデータセットから事前学習する。C-BoW の埋込み単語ベクトル長は 500、文脈語のウィンドウサイズは前後に 5 単語ずつとして、負例数 20 の負例サンプリングを用いて学習した。これより、CoMNet のエンコーダとデコーダの発話文単語列埋込み層、及びデコーダの単語列生成層の語彙ユニット数は 6554、埋込みユニット数は 500 である。また、エンコーダ LSTM とデコーダ LSTM の隠れ層ユニット数は 200 である。

実験では、文内フレーズ注意と会話コンテキスト注意の効果を評価する。注意に基づく応答推論の頑強性の評価のために、データセットの各質問文に対して、文頭と文末、及び文中のランダムな 1 箇所計 3 箇所に未知語を表す ' $\langle unk \rangle$ ' を挿入したテスト用データセットを作成した。応答推論性能評価指標として、入力発話文に対して生成される応答発話文 $s_k (s_k \in S)$ と与えられた教師応答文集合 $E_k = \{e_k\} \in E^*$ の整合性を n-gram 一致度に基づいて測る式 (5) の指標を用いる。いま、n-gram x について、 $m(s_k, x)$, $m(e_k, x)$ をそれぞれ x が s_k , e_k に現れる回数とする。また、 $g(n, s_k)$, $(n = 1, 2, 3)$ を s_k に含まれる n-gram の数とする。このとき、

$$C(S, E^*) = \prod_{n=1}^3 \left(\frac{\sum_k \rho(s_k, E_k) c(n, s_k, E_k) + \delta_n}{\sum_k g(n, s_k) + \delta_n} \right)^{1/3} \quad (5)$$

$$c(n, s_k, E_k) = \sum_x \max_{e_k \in E_k} \min(m(s_k, x), m(e_k, x))$$

$$\rho(s_k, E_k) = \begin{cases} 1 & \text{if } |s_k| \geq \min_{E_k} |e_k| \\ e^{-\frac{\min_{E_k} |e_k|}{|s_k|}} & \text{otherwise} \end{cases}$$

を適合度と呼ぶ。 $C(S, E^*)$ は、1-gram、2-gram、3-gram の適合度の相乗平均である。ここで、 $\rho(s_k, E_k)$ は生成文の短さ

ペナルティ, δ_n は修正係数で, $n = 1$ のとき 0 , $n > 1$ のとき小さな正数である. また, $|s_k|$, $|e_k|$ は単語数である. 教師応答文を集合とするのは, 同じ入力発話文に異なる教師応答文が対応することがあるためである. 学習は約 300 エポックで収束状態に到達した.

文内フレーズ注意のみ組み込んだ場合と注意を用いない場合の, 学習で用いたデータセット, 及びテストデータセットに対する適合度を表 1 に示す. 文内フレーズ注意を組み込む場合は, エンコーダ LSTM を双方向または単方向にした場合の比較も示す. 文内フレーズ注意により頑強な応答推論が実現されることが確認された. 一方, 入力発話文の双方向エンコーディングは性能向上に寄与しなかった.

表 1: 文内フレーズ注意の効果 (適合度)

CoMNet のタイプ	学習データ	テストデータ
注意なし	0.964	0.0676
フレーズ注意 (単方向 LSTM)	0.964	0.813
フレーズ注意 (双方向 LSTM)	0.947	0.775

次に, 会話コンテキスト注意を組み込んだ場合の適合度を表 2 に示す. 発話履歴キュー長は 1, 3, 5 のいずれかに設定している. 発話履歴キュー長が 1 のケースは, 入力発話単語列の単語ベクトル和を追加的に応答生成に用いているだけである. また, エンコーダは単方向 LSTM を用いている. 会話コンテキスト注意, 及びそれを文内フレーズ注意と組み合わせることに伴う頑強な応答推論への効果が確認された.

表 2: 会話コンテキスト注意の効果 (適合度)

CoMNet のタイプ	学習データ	テストデータ
コンテキスト注意のみ (履歴長 1)	0.960	0.163
コンテキスト注意のみ (履歴長 3)	0.970	0.141
コンテキスト注意のみ (履歴長 5)	0.964	0.138
コンテキスト&フレーズ注意 (履歴長 1)	0.962	0.839
コンテキスト&フレーズ注意 (履歴長 3)	0.961	0.849
コンテキスト&フレーズ注意 (履歴長 5)	0.959	0.854

5.3 人物再同定実験

ソフトウェア演習での運用で収集した画像データを用いた人物識別性能評価と, 映像データを用いた人物再同定性能評価を行った. 人物識別性能評価では, 30 人分の上半身矩形画像 6000 枚と誤検出分類のための非人物矩形画像 200 枚の計 6200 枚の画像からなるデータセットを用意した. このうち, 学習には 4960 枚を用いて 5 分割交差検証を行い, テストには 1240 枚を用いた. ランダムフォレストの主なパラメータは, 決定木数 100, 決定木の最大深さ 15, サブサンプリング率 0.5 である. テストデータの 31 クラスに対する正解率は 0.84 であった.

人物再同定性能評価では, 移動ロボットが教室内を巡回中にキャプチャした映像を用いる. そして, 映像から人物矩形を検出し, 特徴抽出とインタラクティブなラベル付けを行って学習データを生成し, それまでの学習データからのランダムフォレスト分類器の学習と 1 つ前に作成した分類器を用いた人物識別を同時に一定時間間隔で繰返し行う. 人物矩形検出には, 人物上半身矩形画像 4690 枚と背景画像 8098 枚の計 12788 枚から事前に学習した検出器を使用した. ランダムフォレストのパラメータは人物識別評価の場合と同じである. 時間に沿って順次生成・更新される 7 つの分類器のクラス数と正解率の推移を表 3 に示す. ここで, クラス数は (人物の数+1) である. 序盤は学習データが少ないこととそれまでの学習データに含まれない学生が増えてくることから正解率は低いが, 時間経過と

ともに多くの学生のような見えのデータが収集されて学習に利用されることで正解率は高くなっていく.

表 3: 人物再同定性能評価

時区間	1	2	3	4	5	6	7
クラス数	12	24	28	33	37	38	40
正解率	0.23	0.12	0.15	0.42	0.54	0.44	0.63

6. 考察: CoMNet の関連研究

CoMNet の会話コンテキスト注意に関連する研究として, [Bordes 16] は会話履歴と質問から注意コンテキストを抽出してそれに最も適合する応答文を選択しているが, 応答文は候補集合からの選択に限定される. 一方, CoMNet では注意フレーズと注意コンテキストを融合して応答文を生成している点, 及び注意コンテキストを発話の順序と影響度の発話履歴キューによる管理のもとで計算している点の特徴である.

7. むすび

本論では, オープンチャットとロボットの関係によりティーチングアシスタントと協働して演習授業支援をする統合システムに関して, システム構成とその要素技術としての発言分類, CoMNet による応答自動生成, 人物再同定の方法, 及び運用で収集したデータを用いた実験による有効性評価について述べた. 今後, チャットボットに関しては, 主に話題カテゴリ「エラー・トラブル」, 期待応答タイプ「原因」「方法」における外部知識参照を伴う文脈依存対話型応答生成, 移動ロボットの人物識別・再同定に関しては, 学生の学習状態推定, 及び外部学習履歴 DB との学習支援のための関係を, 従来同様, 授業支援運用をフィードバックしながら開発し統合していく.

参考文献

- [Maderer 16] Maderer, J.: Artificial Intelligence Course Creates AI Teaching Assistant, <http://www.news.gatech.edu/2016/05/09/artificial-intelligence-course-creates-ai-teaching-assistant> (2016)
- [Mubin 13] Mubin, O., Stevens, C.J., et al.: A Review of the Applicability of Robots in Education, *J. of Technology for Education and Learning*, Vol.1, pp.1-7 (2013)
- [渥美 17] 渥美雅保, 村田祐樹, 安川葵: SOTARO: オープンチャットとロボットの関係による演習授業支援システム - Human-in-the-loop 型機械学習によるアプローチ -, 情報処理学会第 79 回全国大会, 6E-07 (2017)
- [村田 17a] 村田祐樹, 渥美雅保: 移動ロボットによる人物識別のためのインタラクティブな教師付き機械学習とその演習授業への適用, 情報処理学会第 79 回全国大会, 4ZC-08 (2017)
- [村田 17b] 村田祐樹, 渥美雅保: Random Forest と畳み込み特徴に基づくインタラクティブ機械学習による人物識別, 2017 年度人工知能学会全国大会, 4K1-3 (第 31 回) 論文集 (2017)
- [安川 17] 安川葵, 渥美雅保: ソフトウェア演習支援のためのチャットシステムへの機械学習の組み込みと授業での運用実験, 情報処理学会第 79 回全国大会, 4ZC-01 (2017)
- [Graves 05] Graves, A. and Schmidhuber, J.: Framewise Phoneme Classification with Bidirectional LSTM Networks, *IJCNN* (2005)
- [Gers 00] Gers, F.A., Schmidhuber, J. and Cummins, F.: Learning to Forget: Continual Prediction with LSTM, *Neural Computation*, Vol.12, No.10, pp.2451-2471 (2000)
- [Bahdanau 15] Bahdanau, D., Cho, K. and Bengio, Y.: Neural Machine Translation by Jointly Learning to Align and Translate, *ICLR* (2015)
- [Bordes 16] Bordes, A. and Weston, J.: Learning End-to-End Goal-Oriented Dialog, arXiv:1605.07683 (2016)