Ontology-based Natural Language Processing for In-store Shopping Situations

Sabine Janzen
Furtwangen University
Furtwangen, Germany
sabine.janzen@hs-furtwangen.de

Wolfgang Maass
Furtwangen University
Furtwangen, Germany
wolfgang.maass@hs-furtwangen.de

Abstract—Natural Language communication between customers and products within in-store shopping environments enables new forms of product interfaces and an improved filtering and intuitive presentation of product information. In this article, we describe how customer’s access to product information at the point of sale can be improved through the use of dialogue systems and heterogeneous web-based representations of product information based on formal ontologies within in-store shopping environments. After considering specific requirements of in-store shopping environments on dialogue systems, we present the model of a Conversational Recommendation Agent (CoRA), a domain-specific dialogue system, which realizes an ontology-based Natural Language Processing system for shopping situations.

Keywords—Dialogue system; Natural Language Processing; Ontology; Web-based representations

I. INTRODUCTION

Today, customers in brick-and-mortar stores often lack access to huge amounts of helpful product information on the Web, including product manuals, consumer and professional reviews, and feature comparisons. On the other hand, customers who have in-store Web access through a mobile device may suffer from information overload due to the sheer quantity of product information available. We investigate how customer’s access to product information can be improved through the use of product-centered dialogue systems within in-store shopping environments. The direct embedding of communication services into a product [1] forms the basis for communication between users and products in shopping situations. Natural language based communication within a physical shopping environment enables an improved filtering and presenting of product information [2]. In order to realize such natural language communication between customers and products in purchase situations, dialogue systems and appropriate web-based knowledge representations are necessary. In this contribution, we focus on the following issues: the design of the Natural Language Processing (NLP) approach concerning perspectives of discourse planning [3] and the processing of diverse heterogeneous information resources that appear in purchase situations. We introduce the concept of an ontology-based dialogue system for consumer-product communication used in in-store shopping environments. Next we will illustrate our motivation by an example. Section 3 describes requirements for dialogue systems that support in-store shopping communication between consumers and vendors. Thereafter a dialogue system is described that meets identified requirements (Section 4). Within Section 5, this approach is illustrated by an example. Section 6 presents the results of our user study to test the approach. In Section 7 related work is discussed, followed by a summary of results and future work (Section 8).

II. AN EXAMPLE

The motivation of dialogue systems for in-store shopping situations shall be described by a futuristic example. A customer enters a fashion store because she has seen a nice red jacket in a current film. She carries a smart phone that she generally uses for getting additional and personalized information about products. In the basement she finds this jacket in yellow. Her mobile device receives the product’s ID via RFID-based communication. She asks: “Which jackets are available in red?” The Jacket answers via mobile device: “There are 3 jackets available in red: Cosmic, Flower and Fleur. You can find it on the second floor. If you like, somebody will get it for you.” Customer: “Yes, please.” Jacket: “Are you interested in fashion videos with jacket Cosmic?” Customer: “Why not!” A video with Scarlett Johansson wearing this jacket is shown on a display nearby. Customer: “Very good. Which pair of trousers fit to this jacket?” Jacket: “The best two options are these two trousers.” Both are shown on her mobile device. Jacket: “If you buy this jacket and one of these trousers you will get a 5% discount.” Customer: “That is a good deal.”

III. REQUIREMENTS FOR DIALOGUE INTERACTION WITHIN IN-STORE SHOPPING SITUATIONS

This example indicates some complex relationships between customers, retailers, products and external contents. In heterogeneous shopping environments, information is retrieved from different sources with different semantics. In the following, we have listed key requirements for dialogue systems used in in-store shopping situations that make use of such heterogeneous information sources: (R1) Processing of applied dialogues and related information needs of the customer in shopping situations; (R2) two-dimensional resolution of several, potentially opposing intentions, e.g. of
customers and retailers; (R3) efficient NLP approach for mobile devices as mediators between customers and products; and (R4) processing of heterogeneous data resources, e.g. product information or linguistic knowledge representations.

IV. SYSTEM ARCHITECTURE

A flexible and light-weight dialogue system approach is required in order to process multi-intentional applied dialogues within in-store shopping situations based on heterogeneous resources and mobile devices with limited resources. According to aforementioned requirements (cf. Section 3), we present a model of a domain-specific ontology-based Conversational Recommendation Agent (CoRA) that constitutes a communicative interface between customers and physical products (cf. Fig. 1).

A. General Architecture

The model of the CoRA consists of four basic components: middleware layer Tip ’n Tell [4], mobile client, dialogue webservice and knowledge base (cf. Fig. 1).

The middleware layer is represented by a dedicated middleware for value-added mobile recommendation services, called Tip ’n Tell. Within a purchase situation, Tip ’n Tell requests a pool of product descriptions of logical semantic data structure, called Smart Product Description Object (SPDO) [5], and selects suitable SPDOs and reprocesses them for upcoming requests by CoRA. Currently, we map product data of a leading catalog company in Germany (some 30,000 items) into the SPDO schema. Each physical product is described by one SPDO. Different SPDOs can be stored on distributed servers (R4).

The communication interface is implemented on a mobile client (R3) and displays the results of the purchase conversation, e.g. answers, images etc. in an efficient way. The dialogue webservice is the core component of CoRA. It consists of three modules: Behavioural Agent, Discourse Model and Discourse Management. The Discourse Model represents user intentions (R1) and vendor intentions (R2) as well as the Focus Stack and a discourse history. The Focus Stack represents the course of the past dialogue based on salient concepts. The discourse history covers all dialogue acts concerning the product in focus for building up a product memory stored in a SPDO. The Schema Processing component within the Discourse Management represents question interpretations; the PlanOperator Processing component generates answers according to diverse intentions of participants within different contexts (R2). The knowledge base of CoRA contains knowledge specific to question answering formalized by semantic representations, e.g. schemas, plan operators. Within the next subsections discourse intentions, question schemas, plan operators and ontological representation are described in detail.

B. Discourse Intentions and Question Schemas

The discourse intentions are represented within the Discourse Model component (cf. Section 4.1). They are closely linked with question schemas that are processed within the Discourse Management module and stored in the knowledge base of CoRA. At the beginning, we collected a speech corpus of sales conversations and consulting talks within the product domain of consumer electronics at a trade fair. After transcribing and analyzing the corpus, we derived discourse intentions and question schemas in the context of purchase conversations.

1) Discourse Intentions: After analyzing the text corpus concerning communicative user intentions within the conversation, we created a taxonomy of communicative goals for shopping situations [6]. Besides these user-centered communicative intentions, discourse intentions cover linguistic intentions and vendor-centered intents for participating in a purchase conversation (cf. Section 4.1). Consequently, there are three sorts of discourse intentions: (1) linguistic intentions [7]; (2) user-centered intentions, i.e. communicative intentions of a customer during a discourse and (3) vendor-centered intentions, e.g. satisfaction of a consumer’s information needs and increase of profitability and revenue. Along with communicative intentions of a consumer, vendor-centered intents compose the basis for a model of a two-dimensional goal resolution, i.e. processing of differentiated partly oppositional intentions of several participants within a discourse. In our configuration vendor-centered intentions are subordinated to linguistic and user-centered intentions [8]. Currently, we focus on customer’s communicative intents derived from the corpus. We assume that the main intention of vendors and manufacturers is the increase of profitability and revenue considering contextual changes, e.g. tenses of economic boom or financial crises.

2) Question Schemas: Based on the intentions of the taxonomy of communicative goals, we analyzed the structure of questions within the corpus. Consequently, we identified 19 question schemas that appear frequently within sales
dialogues. The following list shows an extract of identified text schemas:

- **Equation Difference**, e.g. "Whereby trouser A differs from the pair of trousers B?"
- **Description Survey NumberOfProducts**, e.g. "Which navigation systems are available in the color red?" "Which skirts are priced less than € 80?"
- **AdditiveProduct Survey Matching**, e.g. "Which product matches with product A?" "Which colors match with product A?"

Natural Language Understanding (NLU) in CoRA uses a schema-based approach [6]. The application of text schemas within the Natural Language Processing in in-store shopping environments has several reasons and holds diverse advantages, for instance the clear structure of the domain of shopping situations concerning emerging conversation courses [9]. Schemas are appropriate mechanisms to represent these structures and courses of sales conversations. Furthermore, they are useful compilations of stereotyped planning structures and courses of sales conversations. Furthermore, text schemas enable the conjunction of the formalist and functional perspective [3]. They offer the basis for an efficient NLP within physical and mobile environments (cf. Section 3). These requirements refer to two perspectives of discourse planning: functionalist and formalist perspective [3]. We assume that a combination of a functionalist and a viable formalist perspective [3], [11] constitutes an efficient approach for interactive communication with products in in-store shopping environments. Therefore, the approach of the NLG in CoRA is represented by a combination of text planning technologies [12]; rhetorical relations [13], [14], [3]; text schemas [6] and focus of attention [15].

C. Goal-driven Text Plans

After designing the Natural Language Understanding in CoRA, we developed the Natural Language Generation (NLG) approach that primarily meets requirements R1, R2 and R3 (cf. Section 3). These requirements refer to two perspectives of discourse planning: functionalist and formalist perspective [3]. We assume that a combination of a functionalist and a viable formalist perspective [3], [11] constitutes an efficient approach for interactive communication with products in in-store shopping environments. Therefore, the approach of the NLG in CoRA is represented by a combination of the NLG in CoRA and text schemas [6] and focus of attention [15]. The linkage of goal-driven planning technologies with the light-weight constitution of syntactical structures within the schemas enables the conjunction of the formalist and functionalist perspective of discourse planning [3]. Furthermore, the application of planning technologies allows an improved mapping of intentional structures [11] and provides a basis for realization of a two-dimensional goal resolution. The text plans in CoRA are represented by 12 plan operators [12], which are based on an analysis of our corpus and assigned to the 19 text schemas for question understanding. A plan operator defines a text plan for achieving an effect that satisfies a consumer’s communicative intention(s). A further component, the library of rhetorical relations, interlinks parts of a plan operator [13], [14], [3], for instance text schemas or further plan operators. The model of the NLG in CoRA is completed by the focus of attention that represents salient concepts of a discourse [6], [15].

D. Ontological Representation of Intentions, Schemas and Text Plans

The knowledge base of CoRA bases on the Semantic QA Structure Model, which is divided into a domain independent and a domain dependent part (cf. Fig. 2). The Semantic QA Structure Model links concepts of the Natural Language Understanding (NLU) (e.g. text schemas, effect) with concepts of the Natural Language Generation (NLG) (e.g. plan operators, effect). Further, the model integrates a domain-independent lexicon that covers a taxonomy of part-of-speech items. A specific interface concept references lemmas stored in the lexicon for realizing the integration of lexical material into structures of the NLP. Through an instantiation of this domain-independent part of the Semantic QA Structure Model, the domain-dependent layer of the knowledge base is created (cf. Fig. 2). More precisely, the lexicon and NLU/NLG structures are modeled according to the domain (e.g. shopping situation). The non-linguistic part of the knowledge base of CoRA - SPDO - is allocated by the Tip 'n Tell framework and linked with the linguistic structures of the Semantic QA Structure Model.

V. Course of an Exemplary Dialogue Act

Within this section, we want to clarify the coherence between the different modules of CoRA by means of an exemplary dialogue act between a customer and a product within an in-store shopping situation (cf. Fig. 3). Imagine, a customer is situated in an in-store shopping environment and equipped with a RFID capable mobile device (e.g. PDA).

She is interested in a red jacket and wants to know which other jackets are available in red (Fig. 3, step 1). After identification of the product via the mobile device, the CoRA mobile client displays the dialogue interface to the customer. With the aid of the mobile client, the user is able to compose her question based on the selection of single constituents step-by-step. The schema processing module of CoRA processes the question segments selected
by the user, extracts the next possible schema segments from the Semantic QA Structure Model and sends them back to the mobile client (Fig. 3, step 2). The linguistic skeleton of the schema is filled with lemmas of the lexicon and non-linguistic information of the SPDO. CoRA requests the linguistic data of the Semantic QA Structure Model that means traversing the tree of schemas and extracting the next potential segment according to the segment previously selected by the user. When the schema processing module comes across a segment of type SPDO NounPhrase, it fills the slot by requesting Tip ’n Tell for product information, e.g. the product category (jackets). Each question schema consists of question segments in form of linked lists and a linguistic goal amongst others. Every segment points on a part-of-speech in the lexicon and uses constraints for making distinct decisions while traversing the network of schemas. The linguistic goal represents the conversion of a user intention - deduced from the selection of question segments - into a machine processable representation. When a customer completes a question or the Schema Processing component finds that the current segment does not point on another segment, the Question Understanding module detects that the question is completed. It identifies the assigned question schema Description Survey NumberOfProducts (Fig. 3, step 3) and the linguistic goal NumberOf(Product) Res (Jacket.Colour(red)) (Fig. 3, step 4) via requesting the Semantic QA Structure Model. Now, the customer can pose the resulting question “Which jackets are available in red?” (Fig. 3, step 5). In order to answer this question, the linguistic goal is transferred to the effect (KNOW-ABOUT ?agent NumberOf(Product) Res (Jacket.Colour(red))) that fulfills the effect (Fig. 3, step 7). Within the PlanOp Processing module, the text plan of the answer is constructed based on the selected plan operator. Each plan operator consists of a mandatory nucleus-part (e.g. Pres Number Products) and one or more optional satellite-parts (e.g. Matching) referenced through rhetorical relations. The plan operator parts contain schematic structures or further plan operators. Each part of the plan operator is expanded by requesting the Semantic QA Structure Model and the SPDO via Tip ’n Tell. The resulting linguistic material is allocated (Fig. 3, step 8), for instance Pres Number Products: "There are 3 jackets available in red: Cosmic Rain, Flower Beach and Thin Air."; Matching: "The jacket Cosmic Rain fits to the boots Spring." Figure 4 shows an extract of the Semantic QA Structure Model: the ontological representation of the plan operator NUMBER PRODUCTS.
At this point, the vendor intentions are integrated within the optional satellite part Matching (Fig. 3, step 9). Product-comprehensive information, for instance, the product matching or product bundle results, are generated via reasoning based on standardized web-based rules (e.g. SWRL) that are applied to SPDO instantiations of products in Tip ‘n Tell. CoRA requests this information from Tip ‘n Tell and integrates them into the linguistic structure of the expanded plan operator parts. Based on the focus of attention and the course of the conversation, CoRA decides which parts of an expanded plan operator are integrated into the final answer (Fig. 3, step 10). Within the Answer Generation the focus of the satellite Matching is checked against topmost parts of the Focus Stack which represents the course of past discourses based on salient concepts. If a focus already exists within a scope, a satellite will not become a part of the final answer for avoiding repetitions (Fig. 3, step 11). Furthermore focus of plan operator parts that were included in final answers are pushed onto the Focus Stack. Finally, an answer is composed and displayed on the mobile client: “There are 3 jackets available in red: Cosmic Rain, Flower Beach and Thin Air. The jacket Cosmic Rain fits to the boots Spring.” (Fig. 3, step 12) The discourse act consisting of question, answer, effect and timestamp is sent to Tip ‘n Tell for saving a discourse history within the SPDO for building up a product memory (Fig. 3, step 13). On implementation level, CoRA constitutes a module of the Tip ‘n Tell framework and is implemented in form of a Java web service. To realize the handling of the semantic representations, more precisely the SPDO and the Semantic QA Structure Model, we use the Jena Semantic Web Framework2. Furthermore, each part of the knowledge base is formalized by semantic representations (currently in OWL-DL).

VI. USER STUDY

We conducted a user study to test the usability and utility of the interactive dialogue concept of CoRA for in-store product information acquisition. As a first step, the dialog system was therefore simulated on a desktop-based PC. In this sense, we assured that the dialogue concept is understood properly and would make sense to users for in-store product information acquisition in general. A total of 16 subjects were recruited from a German university. First, we presented the interactive dialogue system to each subject. For this demonstration, a purchase scenario within a fashion store was described with Jacket A being the product the questions were directed at. In this sense, the following three questions were constructed step-by-step: a) Which jackets are available in red?, b) Which material is used?, c) Is there a bundle discount available when buying Jacket A together with the scarf 'Flower'? Then, subjects were given a questionnaire to indicate their perceived attitudes towards this interactive dialogue concept. Among the SUS items [16], which resulted in a high usability score of 76.3 (SD: 12.69, Alpha: .81, 10 items), we also used an existing scale from [17] to measure the perceived usefulness of the dialogue concept (Mean: 5.56, SD: .78, Alpha: .80, 4 items). In addition, subjects had to indicate their intention to use the dialogue concept (Mean: 5.56, SD: .78) [18]. Cronbach’s alpha values for all multi-item scales were above .70 as indicated above in brackets and thus show good reliability [19]. Finally, all scores lie highly significant ($p < .001$) over their average values by applying one-sample t-tests, which supports the conclusion that the interactive dialogue concept of CoRA is perceived as both usable and useful for in-store product information acquisition.

VII. RELATED WORK

The interactive communication between customers and products within in-store shopping environments is characterized by a triangular linkage of three fields of research - Ontology Engineering, Natural Language Processing (NLP) and Ambient Intelligence. Through the integration of concepts of NLP into ambient environments, two scientific fields as yet separated are merged and exciting research issues are opened [20]. An AmI system has to deal with large amounts of unstructured information from heterogeneous sources. These have to be selected and prepared to satisfy user information needs [2]. At this point, NLP technologies become relevant, because they enable interpretation, transformation and filtering of information and therefore enhance quality and ease of use of Ambient Intelligence systems. Within NLP, there are four dominant discourse theories [6], [13]–[15]: what they have in common is that discourses are purpose-driven, i.e. communication is driven by user intentions to achieve a communicative goal. Furthermore, there exist two perspectives to consider the planning of discourses: functionalist and formalist perspective [3]. On the other side of the triangle, the linkage of the two research fields NLP and Ontology Engineering facilitates the development of novel QA systems that use ontologies as core knowledge component [8], [21]. We assume, that NLP benefits from the appliance of semantic knowledge representations. Ontologies are important components of information systems and information processing, because they support the interoperability between information systems and avoid isolated applications [22], [23].

VIII. CONCLUSION AND FUTURE WORK

Natural language dialogue interaction between customer and products within in-store shopping environments encompasses the capability for a redesign and extension of communication interfaces between customers, vendors and manufacturers at the point of sale. In this article, we described how customer’s access to targeted product information at the point of sale can be improved through the use of dialogue systems and heterogeneous web-based representations of product information based on formal ontologies.
We introduced with the Conversational Recommendation Agent (CoRA) the model of a dialogue system that realizes an ontology-based NLP between customers and products based on questions and answers. In our future work, we will focus on three issues: (1) a more differentiated processing of dynamically changing intentions of different stakeholders within the generation of answers in real-time, e.g. vendor-centered intention resolution is configurable by vendors, (2) improving the flexibility of answer generation through integration of user models [24] and (3) representation of courses of conversations and those integration into the discourse model of CoRA based on tasks [8].

ACKNOWLEDGMENT

This paper resulted from project "Smart Product Networks" funded by the German Ministry for Education and Research (BMBF).

REFERENCES