Culture-Aware Dialogue Management for Conversational Assistants

Juliana Miehle, Nicolas Wagner, Wolfgang Minker, and Stefan Ultes

Abstract The cultural background has a great influence on the people's behaviour and perception. With the aim of designing a culturally sensitive conversational assistant, we have investigated whether culture-specific parameters may be trained by use of a supervised learning approach. We have used a dialogue management framework based on the concept of probabilistic rules and a multicultural data set to generate a culture-aware dialogue manager which allows communication in accordance with the user's cultural idiosyncrasies. Hence, the system response to a user action varies depending on the user's culture. Our data set contains 258 spoken dialogues from four different European cultures: German, Polish, Spanish and Turkish. For our evaluation, we have trained a culture-specific dialogue domain for each culture. Afterwards, we have compared the probability distributions of the parameters which are responsible for the selection of the next system action. The evaluation results show that culture-specific parameters have been trained and thus represent cultural patterns in the dialogue management decision process.

1 Introduction

We live in a globally mobile society in which people of widely different cultural backgrounds live and work together. The number of people who leave their ancestral cultural environment and move to countries with different culture and language is increasing. This spurs the need for culturally sensitive conversation agents, especially

J. Miehle (✉) · N. Wagner · W. Minker
Institute of Communications Engineering, Ulm University, Ulm, Germany
e-mail: juliana.miehle@uni-ulm.de

N. Wagner
e-mail: nicolas.wagner@uni-ulm.de

W. Minker
e-mail: wolfgang.minker@uni-ulm.de

S. Ultes
Mercedes-Benz Research and Development, Sindelfingen, Germany
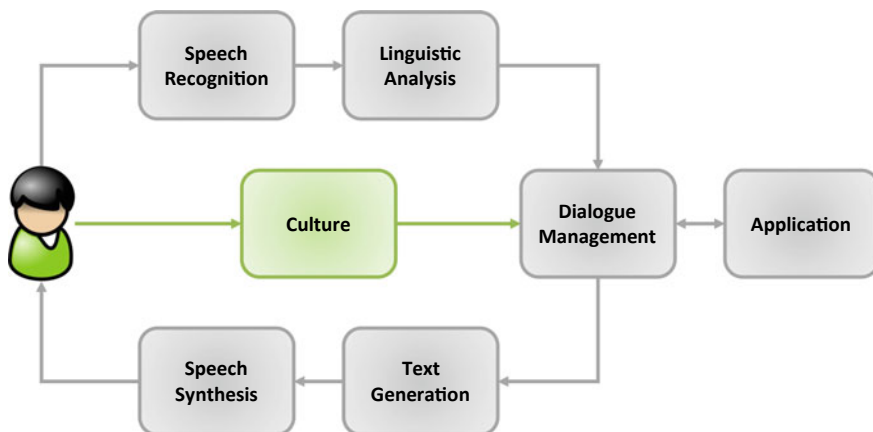e-mail: stefan.ultes@daimler.de

103

Fig. 1 The user's culture is used in the dialogue management to adapt the system behaviour to the user.

for sensitive topics such as health care. Hence, our aim is to design a culture-aware dialogue system which allows a communication in accordance with the user's cultural idiosyncrasies. By adapting the system's behaviour to the user's cultural background, the conversation agent may appear more familiar and trustworthy.

In this work, we use spoken dialogues from four European cultures (German, Polish, Spanish and Turkish) to train a culture-aware dialogue manager (see Fig. 1). The selection of the next system action is thus adapted to the cultural background of the user. We use a dialogue management framework based on the concept of probabilistic rules which combines the benefits of logical and statistical methods to dialogue modelling: The probabilistic rules represent the internal models of the domain in a compact form and the unknown parameters included in the probabilistic rules are automatically estimated from data using supervised learning. We investigate whether it is possible to train culture-specific parameters for these probabilistic rules in order to represent cultural patterns in the dialogue management decision process.

The structure of the paper is as follows: In Sect. 2, we present related work in the field of culture-sensitive interface design. Afterwards, the corpus that has been used in this work is described in Sect. 3 and the design and implementation is outlined in Sect. 4. We present the evaluation and our results in Sect. 5, before concluding in Sect. 6.

2 Related Work

Brejcha [1] has described patterns of language and culture in Human-Computer Interaction and has shown why these patterns matter and how to exploit them to design a better user interface. Reinecke and Bernstein [8] have presented a design approach

for culturally-adaptive user interfaces to enhance the usability. The authors developed a prototype web application that serves as a to-do list tool. During registration, the users have to input their current country of residence, former countries they lived in and the length of stay in each country. Based on this information, the design and content of the web application is adapted to the user.

Furthermore, Traum [9] has outlined how cultural aspects may be included in the design of a visual human-like body and the intelligent cognition driving action of the body of a virtual human. Therefore, different cultural models have been examined and the author points out steps for a fuller model of culture. Georgila and Traum [2] have presented how culture-specific dialogue policies of virtual humans for negotiation and in particular for argumentation and persuasion may be built. A corpus of non-culture specific dialogues is used to build simulated users which are then employed to learn negotiation dialogue policies using Reinforcement Learning. However, only negotiation specific aspects are taken into account while we aim to create an overall culture-sensitive dialogue manager.

A study presented by Miehle et al. [7] investigated cultural differences between Germany and Japan. The results pointed out that communication idiosyncrasies found in Human-Human Interaction may also be observed during Human-Computer Interaction in a Spoken Dialogue System context. Moreover, the study described by Miehle et al. [6] examined five European cultures whose communication styles are much more alike than the German and Japanese communication idiosyncrasies. The results show that there are differences among the cultures and that it depends on the culture whether there are gender differences concerning the user's preference in the system's communication style. These studies show that culture-adaptive behaviour is an important aspect of spoken user interfaces. This is why we address the task of culture-aware dialogue management.

3   Corpus Description

Our data set is based on recordings on health care topics containing spontaneous interactions in dialogue format between two participants: one is acting as the system is expected to perform while the second one is taking the role of the user of the system. Each dialogue is allocated with a unique dialogue ID and each action is assigned

- a dialogue action number,
- a participant,
- a speaker,
- a dialogue action, and
- the original utterance.

The dialogue action number counts from 1 to n starting with the first dialogue action, where n is equal to the number of dialogue actions of the respective dia-

Table 1 The culture distribution of the dialogue actions in the corpus

Culture	#Dialogue actions
German	4849
Polish	1017
Spanish	1002
Turkish	1024

logue. The participant specifies the two roles *system* and *user* and the speaker indicates which of the predefined speakers was talking. Each speaker is identified by an anonymous speaker ID and a separate table contains profile information about each speaker, including the gender, the culture, the age, the country of birth and the current country of residence. The dialogue action is chosen out of a set of 59 distinct dialogue actions which have been predefined in advance. Furthermore, the original utterance (in the original language) is added to each dialogue action and for each utterance, the topics being talked about are annotated (in English). Moreover, for each dialogue, the system's role is specified. The available system roles are defined as *social companion*, *nursing assistant* and *health expert*. Overall, the corpus covers 258 dialogues. The culture distribution of the dialogue actions is shown in Table 1.

4 Design and Implementation

For the implementation of our culture-aware dialogue manager, we have used the open-source software toolkit OpenDial [5]. It combines the benefits of logical and statistical methods to dialogue modelling by adopting a hybrid approach. Probabilistic rules represent the domain model in a structured format and allow system designers to integrate their domain knowledge. These rules contain unknown parameters that can be estimated from dialogue data using supervised learning. Thus, this hybrid concept allows the system designers to integrate domain-dependent constraints into a probabilistic context. The probabilistic rules formalism is described in [4]. Practically, they are defined as *if...then...else* constructs that map logical conditions to a distribution over possible effects. For the action selection, OpenDial provides utility rules that associate utility values to system decisions. They can be used to find the action with the highest expected utility in the current state. Using these utility rules, we have implemented our dialogue domain as described in Sect. 4.1. Afterwards, we have performed parameter estimation as explained in Sect. 4.2.

4.1 Domain Design

For the implementation of our dialogue domain, we have derived the utility rules from the database described in Sect. 3. We have extracted all possible system actions

```xml
<rule id="Greet">
 <case>
   <condition>
     <if var="a_u" value="Greet"/>
   </condition>
   <effect util="theta_Greet0">
     <set var="a_m" value="PersonalGreet+AskMood"/>
   </effect>
   <effect util="theta_Greet1">
     <set var="a_m" value="PersonalGreet+AskTask"/>
   </effect>
   <effect util="theta_Greet2">
     <set var="a_m" value="PersonalGreet+ShareJoy"/>
   </effect>
   <effect util="theta_Greet3">
     <set var="a_m" value="AskMood"/>
   </effect>
   <effect util="theta_Greet4">
     <set var="a_m" value="AskPlans"/>
   </effect>
   <effect util="theta_Greet5">
     <set var="a_m" value="PersonalGreet"/>
   </effect>
   <effect util="theta_Greet6">
     <set var="a_m" value="Greet"/>
   </effect>
   <effect util="theta_Greet7">
     <set var="a_m" value="MorningGreet"/>
   </effect>
 </case>
</rule>
```

Fig. 2 Implementation of the rule for greeting based on the corpus

a_m in response to a user action a_u, regardless of culture. Since three or more consecutive dialogue actions occur rarely in the data, we have limited the number of possible system actions as response to a user action to two. Afterwards, we have implemented a rule for every user action. Overall, we have seven user actions:

- Accept
- Declare
- Goodbye
- Greet
- Reject
- Request
- Thank

As an example, the implementation of the rule for greeting is depicted in Fig. 2. The rule gets activated if the condition is true, i.e. if the user action is *Greet*. Since it is possible to react with one or two consecutive dialogue actions, there are eight possible effects for the next system action. This design approach ensures that only

Fig. 3 Example of a
dialogue transcript based on
the corpus

```
<interaction>
  <userTurn>
    <variable id="a_u">
      <value>Greet</value>
    </variable>
  </userTurn>
  <systemTurn>
    <variable id="a_m">
      <value>PersonalGreet</value>
    </variable>
  </systemTurn>
</interaction>
```

reasonable pairs of dialogue actions that are covered in the database are included in
the domain.

4.2 Parameter Estimation

After implementing the dialogue domain which captures the possible system actions
in response to every user action, we have used the supervised learning approach based
on the so-called Wizard-of-Oz learning provided within the OpenDial toolkit in order
to estimate the parameters (e.g. `theta_Greet1`). This learning approach allows
not only to learn from Wizard-of-Oz experiments, but also from dialogue transcripts.
As our corpus contains dialogue interactions between two participants where one is
taking the role of the system while the other one is taking the role of the user of that
system, thus resembling the situation of Wizard-of-Oz experiments, we have created
transcripts of these dialogues as input for our parameter estimation. An example of
such a transcript is shown in Fig. 3. In this *interaction*, the user action (`a_u`) *Greet*
is followed by the system action (`a_m`) *PersonalGreet*.

In the following, we explain the Wizard-of-Oz learning that has been used for the
parameter estimation. According to Lison [4], a Wizard-of-Oz interaction is defined
as a sequence of state-action pairs

$$\mathcal{D} = \{\langle \mathcal{B}_i, a_i \rangle : 1 \leq i \leq n\}, \tag{1}$$

where $\mathcal{B}_i$ is the dialogue state, a_i the performed wizard action at time i and n the
total number of recorded actions. During the learning process, the goal is to learn the
posterior distribution of the rule parameters θ based on the Wizard-of-Oz training
data set $\mathcal{D}$. The algorithm takes each state-action pair $\langle \mathcal{B}_i, a_i \rangle \in \mathcal{D}$ and updates the
posterior parameter distribution after each pair. This posterior distribution can be
decomposed as

$$P(\boldsymbol{\theta} \mid \mathscr{D}) = \eta \, P(\boldsymbol{\theta}) \prod_{\langle \mathscr{B}_i, a_i \rangle \epsilon \mathscr{D}} P(a_i \mid \mathscr{B}_i; \boldsymbol{\theta}), \qquad (2)$$

where $P(\boldsymbol{\theta})$ is the prior distribution and $P(a_i \mid \mathscr{B}_i; \boldsymbol{\theta})$ represents the likelihood of the wizard selecting the action a_i in the dialogue state $\mathscr{B}_i$ given the rule parameters $\boldsymbol{\theta}$. This likelihood can be expressed as a geometric distribution

$$P(a_i \mid \mathscr{B}_i; \boldsymbol{\theta}) = \eta(1-p)^{x-1} p, \qquad (3)$$

where x is the rank of action a_i in the utility $U(a_i \mid \mathscr{B}_i; \boldsymbol{\theta})$, η is a normalisation factor and p represents the learning rate of the estimation process. For our experiments, we used $p = 0.2$.

As prior distribution $P(\boldsymbol{\theta})$, we selected a Gaussian distribution with a mean value of 5 and a variance of 1. However, the probabilistic model in Eq. 2 contains both continuous and discrete random variables. This leads to a nontrivial inference problem. OpenDial offers a sampling technique called *likelihood weighting* to approximate the inference process to solve this issue [3]. Hence, the posterior parameter distribution is sampled after the likelihood of the wizard action is calculated. The outcome is then expressed as a *Kernel density estimator* which subsequently can be converted into a Gaussian distribution. This procedure is performed as long as training data is available.

5 Evaluation

After implementing the dialogue domain and creating the dialogue transcript files for each culture (German, Polish, Spanish and Turkish) based on the data set described in Sect. 3, we have used the transcript files to train the rule parameters $\boldsymbol{\theta}$ of the dialogue domain for the different cultures. Thus, four different culture-specific domains have been trained. Proceeding from the initial probability distribution (Gaussian distribution, $\mu = 5, \sigma^2 = 1$), each parameter has been trained based on the appearance of the corresponding system action in the data set. Since the parameters are updated after each *user action - system action* tuple, a more frequent occurrence of a system action in the database causes the shifting of the mean value to a higher value. In contrast, a rare occurrence correlates with a lower mean value, reducing the probability that such a system action is selected. This effect is illustrated in Fig. 4. In the following, we will evaluate whether the trained parameters vary among the different cultures and therefore represent cultural patterns.

In a first step, we have evaluated whether 1000 dialogue actions are enough for training the rule parameters as the Polish, Spanish and Turkish data sets each contain slightly more than 1000 dialogue actions (see Table 1). In order to do so, we have split the German data set consisting of 4849 dialogue actions into four subsets, each containing 1000 dialogue actions. After training with each of the four German

Fig. 4 Probability
distribution of each
parameter before training
(blue) and example
probability distributions of
two parameters after
training, representing a
frequently occurring system
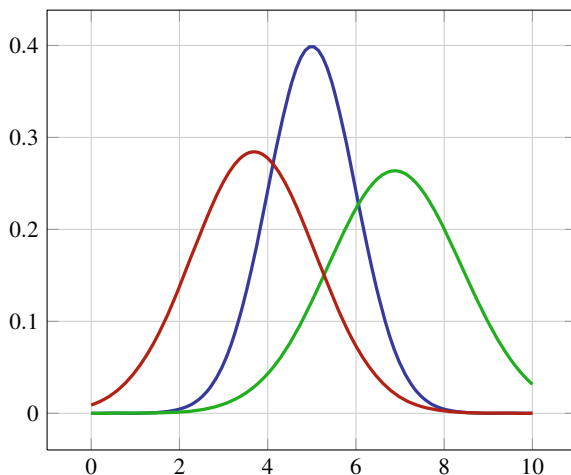action (green) and a rarely
occurring system action (red)

Table 2 Mean values μ and variances σ^2 of the probability distributions of the parameters $\boldsymbol{\theta}$ after training with 1000 German dialogue actions each

Rule	Parameter	German1 $\mu\ (\sigma^2)$	German2 $\mu\ (\sigma^2)$	German3 $\mu\ (\sigma^2)$	German4 $\mu\ (\sigma^2)$
Accept	$\theta_{Request}$	5.20 (2.47)	5.29 (2.84)	5.31 (1.93)	5.43 (1.31)
	$\theta_{Acknowledge+Declare}$	5.27 (2.29)	5.01 (2.78)	4.77 (1.97)	5.00 (2.10)
	$\theta_{Acknowledge+Request}$	5.38 (1.94)	5.18 (3.16)	5.04 (1.87)	4.91 (1.60)
Declare	$\theta_{Request}$	6.88 (52.30)	5.54 (35.26)	7.30 (46.18)	5.12 (12.63)
	$\theta_{Acknowledge+Declare}$	3.69 (48.83)	4.28 (28.63)	5.76 (17.71)	3.69 (39.46)
	$\theta_{Acknowledge+Request}$	9.00 (39.67)	7.38 (24.79)	9.34 (59.98)	6.52 (24.59)
Goodbye	θ_{Thank}	4.95 (1.53)	4.79 (2.10)	4.82 (1.66)	4.76 (2.12)
	$\theta_{SimpleGoodbye}$	5.41 (1.52)	5.97 (2.08)	5.71 (1.52)	5.85 (2.15)
	$\theta_{PersonalGoodbye}$	5.05 (1.23)	5.15 (2.06)	4.63 (2.07)	4.98 (2.04)
Greet	$\theta_{PersonalGreet}$	5.24 (3.24)	4.94 (2.83)	5.52 (3.91)	5.62 (1.87)
	$\theta_{Request}$	5.23 (3.17)	5.54 (1.31)	5.11 (3.20)	5.00 (2.16)
	$\theta_{PersonalGreet+AskTask}$	5.21 (2.30)	4.87 (2.26)	5.56 (2.11)	5.32 (1.94)
Reject	$\theta_{Acknowledge}$	4.99 (1.07)	5.09 (1.10)	5.12 (1.00)	5.08 (1.11)
	$\theta_{Request}$	4.87 (1.06)	5.25 (1.07)	5.00 (1.07)	5.05 (1.07)
	$\theta_{Declare+Advise}$	4.97 (1.19)	4.90 (1.04)	4.97 (0.89)	4.98 (1.06)
Request	$\theta_{Declare}$	6.48 (2.35)	5.44 (2.43)	5.57 (4.30)	4.78 (7.85)
	$\theta_{Accept+Declare}$	6.03 (1.82)	5.29 (2.83)	5.54 (2.37)	6.74 (4.79)
	$\theta_{Accept+Request}$	4.94 (3.21)	5.21 (2.55)	4.83 (4.15)	5.47 (14.70)
Thank	$\theta_{AnswerThank}$	5.17 (1.43)	5.27 (1.38)	5.27 (1.25)	4.96 (1.42)
	$\theta_{Goodbye}$	5.11 (1.38)	5.03 (1.48)	5.15 (1.49)	5.13 (1.33)
	$\theta_{PersonalAnswerThank+Goodbye}$	5.17 (1.77)	5.00 (1.35)	5.21 (1.16)	5.02 (1.48)

Table 3 Mean values μ and variances σ^2 of the probability distributions of the three most highly ranked parameters θ after training with the culture-specific data sets

Rule	German	Polish	Spanish	Turkish
	$\mu\ (\sigma^2)$	$\mu\ (\sigma^2)$	$\mu\ (\sigma^2)$	$\mu\ (\sigma^2)$
Accept	$\theta_{Acknowledge+Request}$ 5.38 (1.94)	$\theta_{Request}$ 5.04 (0.92)	$\theta_{Request}$ 5.13 (1.21)	$\theta_{Request}$ 6.06 (2.80)
	$\theta_{Acknowledge+Declare}$ 5.27 (2.29)	$\theta_{Declare}$ 5.03 (1.02)	$\theta_{Acknowledge+Declare}$ 5.04 (1.24)	$\theta_{Acknowledge+Request}$ 5.38 (3.69)
	$\theta_{Request}$ 5.20 (2.47)	$\theta_{ReadNewspaper}$ 4.99 (1.00)	$\theta_{Acknowledge+Request}$ 4.98 (1.20)	$\theta_{ReadNewspaper}$ 4.88 (3.29)
Declare	$\theta_{Acknowledge+Request}$ 9.00 (39.7)	$\theta_{Request}$ 5.33 (3.02)	$\theta_{Request}$ 5.65 (4.53)	$\theta_{Acknowledge+Request}$ 14.20 (137.18)
	$\theta_{Request}$ 6.88 (52.30)	$\theta_{Acknowledge+Declare}$ 4.98 (4.06)	$\theta_{Acknowledge+Request}$ 5.18 (6.14)	$\theta_{Acknowledge+Declare}$ 3.87 (50.71)
	$\theta_{Acknowledge+Declare}$ 3.69 (48.83)	$\theta_{Acknowledge+Advise}$ 4.94 (4.01)	$\theta_{Acknowledge+Declare}$ 5.17 (4.41)	θ_{Advise} 3.86 (88.75)
Goodbye	$\theta_{SimpleGoodbye}$ 5.41 (1.52)	$\theta_{SimpleGoodbye}$ 5.03 (0.96)	$\theta_{MeetAgainGoodbye}$ 5.34 (1.61)	$\theta_{SimpleGoodbye}$ 5.21 (1.82)
	$\theta_{PersonalGoodbye}$ 5.05 (1.23)	θ_{Thank} 5.03 (1.02)	$\theta_{SimpleGoodbye}$ 5.10 (1.76)	$\theta_{MeetAgainGoodbye}$ 5.00 (1.85)
	θ_{Thank} 4.95 (1.53)	$\theta_{AnswerThank}$ 5.02 (1.03)	$\theta_{Acknowledge}$ 4.98 (2.97)	$\theta_{MorningGoodbye}$ 4.99 (2.36)
Greet	$\theta_{PersonalGreet}$ 5.24 (3.24)	$\theta_{PersonalGreet}$ 5.50 (2.17)	θ_{Greet} 6.21 (2.47)	$\theta_{PersonalGreet+AskMood}$ 5.32 (1.51)
	$\theta_{Request}$ 5.23 (3.17)	$\theta_{PersonalGreet+AskMood}$ 5.34 (3.34)	$\theta_{Declare}$ 5.35 (4.42)	$\theta_{PersonalGreet+AskTask}$ 5.26 (1.97)
	$\theta_{PersonalGreet+AskTask}$ 5.21 (2.30)	θ_{Greet} 5.22 (3.28)	$\theta_{AskMood}$ 5.04 (2.54)	$\theta_{Declare}$ 5.16 (1.74)
Reject	$\theta_{Acknowledge+Declare}$ 5.06 (1.16)	no data available	$\theta_{Acknowledge+Advise}$ 5.11 (1.07)	$\theta_{Request}$ 5.15 (1.20)
	$\theta_{Request}$ 5.04 (1.13)	no data available	$\theta_{Declare+Advise}$ 5.02 (1.59)	θ_{Advise} 5.03 (1.29)
	$\theta_{Acknowledge+Request}$ 5.01 (1.05)	no data available	$\theta_{Request}$ 5.00 (1.73)	$\theta_{Acknowledge+Advise}$ 4.98 (1.11)
Request	$\theta_{Declare}$ 6.48 (2.35)	$\theta_{Reject+Declare}$ 137.04 (40041.25)	$\theta_{Advise+Request}$ 15.51 (187.93)	$\theta_{Declare}$ 6.46 (7.11)
	θ_{Accept} 6.11 (1.89)	θ_{Reject} 129.99 (98009.71)	$\theta_{Reject+Advise}$ 13.10 (373.54)	$\theta_{Accept+Declare}$ 6.44 (5.49)
	$\theta_{Accept+Declare}$ 6.03 (1.82)	$\theta_{Accept+Declare}$ 107.23 (33122.08)	$\theta_{Declare+Request}$ 12.75 (246.97)	$\theta_{Motivate}$ 6.11 (5.07)

(continued)

Table 3 (continued)

Rule	German	Polish	Spanish	Turkish
	$\mu\ (\sigma^2)$	$\mu\ (\sigma^2)$	$\mu\ (\sigma^2)$	$\mu\ (\sigma^2)$
	$\theta_{AnswerThank}$	$\theta_{Acknowledge+Goodbye}$	$\theta_{Goodbye}$	$\theta_{PAnswerThank+Goodbye}$
	5.17 (1.43)	5.02 (0.98)	5.64 (1.27)	5.10 (1.53)
Thank	$\theta_{PAnswerThank+Goodbye}$	$\theta_{Goodbye}$	$\theta_{Acknowledge+Answer|Thank}$	$\theta_{Goodbye}$
	5.17 (1.77)	5.02 (1.01)	5.08 (1.30)	4.98 (1.57)
	$\theta_{Goodbye}$	$\theta_{PAnswerThank+Goodbye}$	$\theta_{Motivate}$	$\theta_{AnswerThank+Advise}$
	5.11 (1.38)	5.02 (0.95)	5.07 (1.34)	4.98 (1.62)

training sets, we have obtained similar alternatives for the action selection as the three parameters with the highest mean values are the same for each rule. The corresponding values for the mean μ and the variance σ^2 of the probability distributions for each rule are shown in Table 2. For some rules (e.g. Declare), we get the same ranking for every subset, for others (e.g. Accept) the mean values of the three most highly ranked parameters differ only slightly. However, each of the four German domains results in a similar system strategy, showing that the relevant information is contained in the data. This allows the assumption that training with 1000 dialogue actions is sufficient to train culture-specific parameters. Furthermore, as the average over all four subsets corresponds approximately to the values of subset *German1*, this subset is used to represent the German culture in the further course of the paper. Moreover, this first part of evaluation has revealed that the variance σ^2 of the probability distributions of the parameters is of little importance in the applied sampling technique. The action selection is mainly based on the mean value μ. Hence, we have based the cross-cultural comparison in the second part of our evaluation on the means.

In the second step, we have used the German, Polish, Spanish and Turkish data sets (each containing approximately 1000 dialogue actions) to train the rule parameters $\boldsymbol{\theta}$ of four culture-specific dialogue domains. The three parameters with the highest mean value for each culture are shown in Table 3. It can be seen that the different characteristics of the cultures occasionally have led to different parameters with highest mean values. In the following, we discuss the similarities and differences for each rule.

Accept

If the last user action is accepting what the system has said, the system either reacts with a request for more information or with giving some information to the user (*Declare* or *ReadNewspaper*, what is a special form of giving information). This is the case for every culture tested in our scenario. However, the difference between the cultures lies in whether the system adds an *Acknowledge* (e.g. "Okay.") or not. While this is very likely for German, it is more unlikely for Polish, Spanish and Turkish (as the parameter with the highest mean does not include it for these cultures).

Declare

If the last user action is a *Declare*, the system may request for more information or give some information to the user. This applies again to each of our cultures. However, we are able to observe two differences between the cultures, namely (1) whether an *Acknowledge* is added or not, and (2) whether the information is presented as an *Advise* or a *Declare*. While it is very likely to add an *Acknowledge* for German and Turkish, it is more unlikely for Polish and Spanish. Moreover, the information will always be presented as a *Declare* (e.g. "He needs help getting up.") for German and Spanish, while an *Advise* (e.g. "You should help him up.") may be used for Polish and Turkish.

Goodbye

After the user says goodbye, the system usually answers with any form of saying goodbye. However, the form differs between the examined cultures. While for German, Polish and Turkish a *SimpleGoodbye* (e.g. "Good bye.") is most probable, for Spanish it is more likely that a *MeetAgainGoodbye* (e.g. "See you.") is used. Moreover, for German it is also very common that the user's name is used in a *PersonalGoodbye* (e.g. "Bye Anna.") and for Polish a *Thank* (e.g. "Thank you.") or an *AnswerThank* (e.g. "You're welcome.") instead of a *Goodbye* might be used.

Greet

If the user greets the system, the most likely system response is also a greet. German, Polish and Turkish uses a *PersonalGreet* (e.g. "Hello Anna."). However, in our Spanish model, the user is not addressed by name but simply welcomed with a *Greet* (e.g. "Hello."). Furthermore, in the German, Polish and Turkish model, an *AskTask* (e.g. "How can I help you?") or an *AskMood* (e.g. "How are you?") are very likely to be added to the greeting. However, for Spanish only one dialogue action is used.

Reject

In our Polish data, it is never the case that the user rejects anything of the system, even if the topics of conversation are evenly distributed among the cultures. For the other cultures, the system usually reacts with a request for more information or with giving some information to the user. For the latter, the difference between the cultures lies in whether the system uses a *Declare* or an *Advise* to present the information: German uses a *Declare*, while Spanish and Turkish use an *Advise*. Moreover, for German and Spanish usually an *Acknowledge* is added.

Request

A request is the most likely user action, what can be seen from the high mean values in this rule. Obviously, answering the user's question and thus giving the requested information is the most likely system response for every culture. However, there are differences in how the information is presented. While in the German, Polish and Turkish model, a *Declare* with optional addition of an *Accept* or *Reject* is used, the Spanish model rather utilises an *Advise* than a *Declare*. Moreover, the Turkish model also uses a *Motivate* (e.g. "Good idea!").

Thank

If the last user action is a *Thank*, the system may react to it with an *AnswerThank*

or interpret it in the way that the user wants to end the dialogue and thus answer with a *Goodbye* (or combine both dialogue actions). As the mean values of the three most highly ranked parameters differ only slightly for every culture, we compare the amount of occurrences of these options among the most highly ranked parameters for each culture. We can see that the Polish model always answers with a *Goodbye*, while the others do not always interpret the user's *Thank* in this way. Moreover, the Spanish model also uses a *Motivate* and the Turkish model may add an *Advise*.

6   Conclusion and Future Directions

With the aim of designing a culturally sensitive conversational assistant, in this work we have investigated whether culture-specific parameters may be trained by use of a supervised learning approach. In order to do so, we have used spoken dialogues from four European cultures, namely German, Polish, Spanish, and Turkish, to train a culture-aware dialogue manager. For the implementation we have used the open-source software toolkit OpenDial [5] which is based on the concept of probabilistic rules. Thus, it combines the benefits of logical and statistical methods to dialogue modelling.

For our evaluation we have trained four different culture-specific dialogue domains. For each culture, we have used a data set containing approximately 1000 dialogue actions as we have shown that 1000 dialogue actions are enough for training the rule parameters. Afterwards, we have compared the probability distributions of the trained parameters. Each parameter is expressed as a Gaussian distribution. Thus, we have examined the differences between the cultures in terms of the mean values of the corresponding probability distributions. The evaluation results have shown that the different characteristics of the cultures result in different parameters with highest mean values. Hence, the system response to a user action varies depending on the culture.

In future work, we will examine whether the proposed approach can be extended to other conversational topics and further cultures. In particular, we are interested in non-European cultures since the differences to the cultures studied in this work might be more significant. Moreover, we plan to conduct an evaluation with real users to see, how a varying action selection based on the culture is perceived.

References

1. Brejcha J (2015) Cross-cultural human-computer interaction and user experiencedesign: a semiotic perspective. CRC Press
2. Georgila K, Traum D (2011) Learning culture-specific dialogue models from non culture-specific data. In: Universal access in human-computer interaction. users diversity. Springer, pp 440–449
3. Lison P (2013) Structured probabilistic modelling for dialogue management. Doctoral dissertation, University of Oslo
4. Lison P (2015) A hybrid approach to dialogue management based on probabilistic rules. Comput Speech Lang 34(1):232–255
5. Lison P, Kennington C (2016) Opendial: a toolkit for developing spoken dialogue systems with probabilistic rules. In: Proceedings of the 54th annual meeting of the association for computational linguistics (Demonstrations). Association for Computational Linguistics, Berlin, pp 67–72
6. Miehle J, Minker W, Ultes S (2018) What causes the differences in communication styles? a multicultural study on directness and elaborateness. In: Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018). European Language Resources Association (ELRA)
7. Miehle J, Yoshino K, Pragst L, Ultes S, Nakamura S, Minker W (2016) Cultural communication idiosyncrasies in human-computer interaction. In: 17th annual meeting of the special interest group on discourse and dialogue, p 74 (2016)
8. Reinecke K, Bernstein A (2013) Knowing what a user likes: a design science approach to interfaces that automatically adapt to culture. MIS Q 37(2):427–453
9. Traum D (2009) Models of culture for virtual human conversation. In: Universal Access in Human-Computer Interaction. Applications and Services, pp 434–440 (2009)