Cultural User Modeling With CUMO: An Approach to Overcome the Personalization Bootstrapping Problem

Katharina Reinecke, Gerald Reif, and Abraham Bernstein

Department of Informatics, University of Zurich
{reinecke,reif,bernstein}@ifi.uzh.ch

Abstract. The increasing interest in personalizable applications for heterogeneous user populations has heightened the need for a more efficient acquisition of start-up information about the user. We argue that the user's cultural background is suitable for predicting various adaptation preferences at once. With these as a basis, we can accelerate the initial acquisition process. The paper presents an approach to factoring culture into user models. We introduce the cultural user model ontology CUMO, describing how and to which extend it can accurately represent the user's cultural background. Furthermore, we outline its use as a re-usable and shared knowledge base in a personalization process, before presenting a plan of our future work towards cultural personalization.

1 Introduction

Personalization of software for heterogeneous user populations has been proven to increase working efficiency and user satisfaction [1]. Nevertheless, only a handful of applications employ user modelling techniques today. A major reason is the tedious collection of assumptions about the user's preferences when employing the user model for the first time. This drawback, the so-called *bootstrapping problem* [2], is especially severe for systems that are not regularly re-visited, such as in the case of cultural heritage systems. The problem can be mitigated if user models were used across several applications. In addition, we believe that the knowledge about the user's cultural background can rapidly expedite the acquisition process: As many preferences are deeply-rooted in a person's cultural background, culture bundles information about a variety of partialities, such as information density, navigational support, the level of hierarchy in the information presentation, or the learning style [3]. Moreover, it can reveal aspects that unconsciously affect a person's processing of information. In the area of cultural heritage systems, the potential advantage of incorporating cultural background into cross-system personalization can be extended to a holistic representation of content collaboration between institutions.

In order to overcome these problems, we propose an approach to cross-system personalization employing the cultural user model ontology CUMO. The idea is a user model that accompanies the user "wherever he goes", no matter which application or device he uses. With that, we will also contribute to reducing the time needed for costly manual localization of software as described in [4].

In the following, we present the state of the art in related research areas, outlining an approach how existing work can be exploited for the implementation of cultural user modeling. On this basis, we introduce our cultural user model ontology CUMO, which is represented in OWL. The subsequent section shows how the ontology can interact with an application in order to establish knowledge about the user's cultural background. We will close with a discussion of our approach, outlining limitations and further research that still needs to be conducted in the endeavour towards cultural personalization.

2 Related Work

Cultural personalization of content and user interfaces bases on the research fields international usability, localization and culture, user modeling and personalization. While these areas have been well-researched, very little effort has been made to combine these fields in an interdisciplinary approach towards culturally adaptive software. Work in this direction includes [5] in the area of e-learning and [6] in the area of cultural adaptivity in navigation systems. These approaches, however, are not conform with our requirements of interoperability and a reusable cultural user model.

Localization and Cultural Dimensions. Classifying culture has been the focus of study by many researchers including cultural theorists Hofstede and Trompenaar [7, 8]. Their cultural dimensions predicate on differences in values, providing measurable classifications of culture. Researchers from the field of user interface design have factored these classifications into their studies [3, 9], with that trying to measure the influence of each dimension on user interface localization and international usability. Many of these studies also show, which aspects of the user interface, such as its navigational structure or the level of guidance, are especially influenced by which dimension [10, 11]. While many localization strategies have built upon these dimensions, some researchers have challenged whether prescriptive models of culture can be applied to the field of user interface design [12]. Due to the elusive nature of cultural background, it is indeed questionable to assign one culture to one nation. This, in turn, emphasizes the need for cultural adaptivity as targeted in our approach.

User and Domain Modeling in the Semantic Web. Approaches to user modeling have often been criticized as being application-specific and, thus, not applicable for holistic personalization. While the Semantic Web has gained more attention, researchers have promoted cross-system personalization [1] in order to share information about the user's preferences between different systems, applications and even devices. In this regard, ontologies provide the means to specify a common understanding of the user modeling domain. Recently, many research projects have tackled the problem of distributed user modeling by developing such shared taxonomies [13, 14], for example in the area of e-Learning [15, 16] or e-Government personalization [17]. A first approach to a general user model ontology was introduced in [18], integrating universal concepts for the description of persons. None of the ontologies, however, include cultural information into their user models.

3 Cultural User Modelling

The term "culture" does not have a unified definition, but has been modified depending on different research disciplines. Generally, it is described as an elusive phenomenon of values, norms, institutions and artefacts [19]. Grasping this intangible nature of cultural background in a cultural user model ontology requires the definition of influencing concepts. However, it is almost impossible to say where cultural influences end and personal characteristics start. Likewise, it is hard to determine what is in the scope of a cultural user model ontology and which concepts should be rather dealt with in a general user model ontology. Due to this smooth transition, we have aligned our ontology with the general user model ontology described in [18], enabling the integration of both ontologies for a refined user model. In this regard, CUMO adds the cultural component to the general user model ontology.

3.1 Representing the User's Culture in CUMO

The cultural dimensions identified by Hofstede offer the possibility to structure culture according to the five different concepts *Power Distance*, *Individualism vs. Collectivism*, *Masculinity vs. Feminity*, *Uncertainty Avoidance* and *Long-term vs. Short-term Orientation*. Hofstede's work resulted in a table that assigns five scores (for the five dimensions) to each of the 72 countries that were surveyed. With the help of these five-dimensional vectors we can classify the user's culture and assign the result to certain personalization strategies. This rigid one-to-one mapping of a single culture to a whole nation, is, however, unlikely to correctly represent culturally ambiguous people with all their influences. We have countered the stereotypical classification by mapping these influences onto tangible concepts, which represent classes and properties in CUMO (see Fig. 1). These concepts serve as an addition to Hofstede's dimensions.

In CUMO, the central concept is the `Person` class with its subclasses `Female` and `Male`. Whether the gender has an influence on culture is often debated; however, knowing about a person's gender has proven to be useful as a control variable [5]. Datatype properties connect the class `Person` with Hofstede's dimensions. According to Hofstede's cultural score table, the properties are then assigned an integer value.

For refinement purposes, we added further concepts that factor cultural influences into the user model ontology. These include classes and properties that describe the cultural influence through different places of residence. CUMO comprises the user's `birthplace`, the `currentResidence` and `formerResidence`, all having the range `Location`. `Location` is further subdivided into the subclasses `Continent` and `Country`, which contain individuals of all continents as well as of all countries listed in ISO 3166 [20]. Furthermore, datatype properties of the range integer record the months / years of residence of each instance

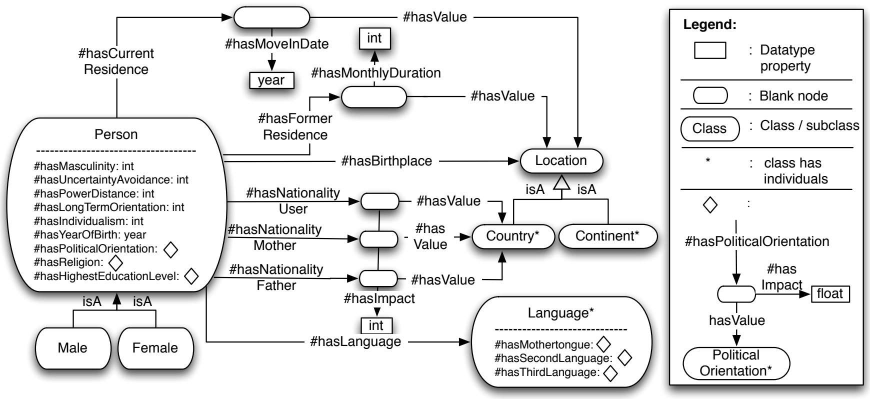

Fig. 1. Culturally Influencing Concepts in CUMO

of `currentResidence` and `formerResidence`. With the help of the datatype property `hasYearOfBirth`, we can therefore roughly calculate the cultural influence of each of these locations on the user. The information taken from the location-related entries cover the fundamentals of cultural influence with regard to cultural values that are firmly anchored in a certain country. Likewise, we deal with the influence of associated persons by including the user's and his parents' nationality. Nationality often stands in for culture, although it is certainly not a synonym [21]. In CUMO, it is therefore only one variable of many.

Going more into detail, CUMO takes into account the user's religion, which is defined as a cultural influence in [22]. The class provides instances of different religious beliefs as well as of major philosophies.

As an additional impact on the user's cultural background according to [7], we have included the class `PoliticalOrientation`, with instances describing different politics. Furthermore, CUMO describes the `HighestEducation- Level`, which acts as a rough indicator for the knowledge about other cultures.

As described in [21], a person's mothertongue also reveals information about his culture, as languages encode cultural information. We have therefore complemented CUMO with the class `Language`, which is divided into the subclasses `Mothertongue`, `SecondLanguage` and `ThirdLanguage`.

The classes `Religion`, `PoliticalOrientation`, `HighestEducationLevel` as well as the three language classes all refer to a datatype property modelling the impact on the user's cultural background. This impact factor can be customized by the application, the user, or both.

3.2 The Modelling Process

For the definition of concepts in CUMO, we assumed to acquire information about the user from both an initial questionnaire and the user's interaction with

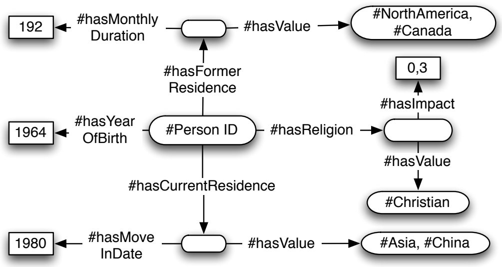

Fig. 2. An Example of a Cultural User Model

an application. The acquisition process is defined by the particular application that is connected to CUMO. Hence, the following section describes an example process as it is conceived for our test application.

The process starts with an initial questionnaire covering the most important entries in the ontology. The more questions the user answers, the more detailed his cultural profile gets; however, the number of questions has to be balanced with the user's willingness to spend time for a personalized profile. In this example, we will therefore introduce the process with a minimum set of questions. The user model is built up according to Fig. 2, which represents an example of a cultural user model.

Due to the correlation to the cultural dimensions that require the information about a country, the first question should ask the user about his current place of residence. The information provided in this first answer allows us to look up the corresponding cultural dimensions in a connected database and pass them on to the cultural user model. With the knowledge about the value of each of the five dimensions, the system can trigger first adaptations. The accuracy of the assigned values for the dimensions can be verified and improved with a next set of questions. If the user provides information about the move in date and about former residences, we can derive the percentage influence by these countries from his indication about the length of stay at each residence. In our example, we have ascertained that the user has lived in China since 1980. To make this information usable, we have acquired about the user's year of birth (1964). Deriving from these facts, we can already assume that he spent 16 years in another country. However, only after the user provides the information that he has lived in Canada for 192 months (or 16 years), we can calculate the influence of both countries. The user model can be more refined with information about the religion. In this case, the user is Christian and the cultural impact is estimated to be 0.3, meaning a 30% influence on the user's culture. Consequently, the system should factor the 30% most obvious religious rules, such as religious colours, into the adaptation process.

In addition to the initial questionnaire, the user interaction with the application can provide us with information about acceptance and accuracy of the proposed adaptation. Tracking mouse movements, such as how much time is spent hovering or how many errors were made, gives information about how the user copes with the adaptation. The next section refers to the future work that is planned in this regard.

4 Limitations and Future Work

Beyond collecting facts about general user preferences, we have included information about the user's cultural background into the user model. With that, we are able to build up a knowledge base that includes the user's origin and culturally influencing factors. A limitation, however, results from the elusive nature of cultural background. In order to gather concepts describing a user's cultural background, we would need to have an exact definition of influences that form a person's culture. This definition would then have to be demarcated from a person's characteristics to be able to restrict the scope of the cultural user model ontology. What, for example, if a user is assigned a single culture, but is culturally influenced by media or foreign friends? This information is not impossible to include, but it is questionable whether users would be willing to provide more information.

Another constraint follows the use of Hofstede's dimensions. Firstly, Hofstede included 72 countries in his investigations. More research is needed to classify the remaining countries. Secondly, Hofstede's classification is not sufficient for a comprehensive representation of a user's cultural background. Adding the information from other influencing factors can certainly help to refine the user profile. However, we need to evaluate whether this model is subtle enough.

At the moment, we are implementing a test application, which will be connected to the user and domain model. The application will be able to receive information from the user model and, after logging the user interaction, pass it back. With that, we plan to evaluate the adaptation process with people from two different cultural backgrounds. Usability tests will then round off the results from our cultural user model and give us information about potential improvements.

5 Conclusion

Findings in research on culture and internationalization indicate that a person's cultural background provides comprehensive information about adaptation preferences. On this account, we have proposed to incorporate culture into a user model in order to overcome the bootstrapping problem in personalization mechanisms. After presenting a way to define a user's cultural background and its influencing factors, we introduced our cultural user model ontology CUMO. The ontology lays the foundation for a rapidly expedited acquisition process, as well as for an automatized localization in general. Used in combination with information systems, it allows to share information about the user's preferences between applications and devices.

References

1. B. Mehta and W. Nejdl. Intelligent Distributed User Modelling: from Semantics to Learning. In *UbiDeUM*, Corfu, Greece, 2007.
2. M. van Kleel and H. Shrobe. A Practical Activity Capture Framework for Personal, Lifetime User Modeling. In *Proc. of the 11th Int. Conf. on UM*, Corfu, GR, 2007.
3. G. Ford and H. Gelderblom. The Effects of Culture on Performance Achieved Through the Use of Human Computer Interaction. In *Proc. of SAICSIT '03*, 2003.
4. K. Reinecke and A. Bernstein. Culturally Adaptive Software: Moving Beyond Internationalization. In *Proc. of the HCII, Beijing, China*. Springer, 2007.
5. E. Kamentz and T. Mandl. Culture and E-Learning: Automatic Detection of a Users' Culture from Survey Data. In *Proc. of IWIPS 2003*, Berlin, Germany, 2003.
6. R. Heimgärtner. Towards Cross-Cultural Adaptive Human-Machine-Interaction in Automotive Navigation Systems. In *Proc. of IWIPS 2005*, Amsterdam, NL, 2005.
7. G. Hofstede. *Culture's Consequences: Comparing values, behaviours and organisations across nations*. Sage Publications Inc., 2003.
8. F. Trompenaars and C. Hampden-Turner. *Riding the Waves of Culture: Understanding Cultural Diversity in Business*. Nicholas Brealey Publishing Ltd, 1997.
9. Aaron Marcus and V.-J. Baumgartner. A Practical Set of Culture Dimensions for Global User-Interface Development. In *Computer Human Interaction*, volume 3101 of *Lecture Notes in Computer Science*, pages 252–261. Springer, 2004.
10. E. W. Gould, N. Zakaria, and S. Yusof. Applying Culture to Website Design: A Comparison of Malaysian and US websites. In *Proceedings of the 18th Annual Conference on Computer Documentation, IEEE*, New Jersey, USA, 2000.
11. A. Marcus. Cross-Cultural User-Interface Design. In *Proc. of the International Conference on Human-Computer Interaction*, volume 2, New Orleans, LA, USA, 2001. Lawrence Erlbaum Associates.
12. A. Smith and Y. Chang. Quantifying Hofstede and Developing Cultural Fingerprints for Website Acceptability. In *Proc. of IWIPS 2003*, Berlin, Germany, 2003.
13. B. Zhou et al. Web Usage Mining for Semantic Web Personalization. In *Workshop on Personalization on the Semantic Web*, pages 66–72, Edinburgh, UK, 2005.
14. F. Abel, I. Brunkhorst, N. Henze, and D. Krause. Personalized Portals – Personal Reader Framework, and Prototypes MyEar and MyNews, 2007. Deliverable, REWERSE project.
15. P. Dolog and W. Nejdl. Challenges and Benefits of the Semantic Web for User Modelling. In *Adaptive Hypermedia 2003*, 2003.
16. L. et al. Aroyo. Interoperability in Personalized Adaptive Learning. *Educational Technology and Society*, 9 (2):4–18, 2006.
17. L. Stojanovic and S. Thomas. Fostering Self-Adaptive e-Government Service Improvement Using Semantic Technologies, 2006. FIT project, `http://www.fit-project.org`.
18. D. Heckmann et al. GUMO - the General User Model Ontology. In *The 10th International Conference on User Modeling*, Edinburgh, UK, 2005. Springer.
19. T. F Hoult. *Dictionary of Modern Sociology*. Littlefield, Adams & Co, Totowa, New Jersey, 1969.
20. International Organization for Standardization. Iso3166. http://www.iso.org.
21. K. Rhoads. *The Culture Variable in the Influence Equation*. Routledge, 2008.
22. D. Jary and J. Jary. *The HarperCollins Dictionary of Sociology*. Collins, 1992.