# Knowledge Representation for Culturally Competent Personal Robots: Requirements, Design Principles, Implementation, and Assessment

Barbara Bruno[1] · Carmine Tommaso Recchiuto[1] · Irena Papadopoulos[2] · Alessandro Saffiotti[3] · Christina Koulouglioti[2] · Roberto Menicatti[1] · Fulvio Mastrogiovanni[1] · Renato Zaccaria[1] · Antonio Sgorbissa[1]

## Abstract

Culture, intended as the set of beliefs, values, ideas, language, norms and customs which compose a person's life, is an essential element to know by any robot for personal assistance. Culture, intended as that person's background, can be an invaluable source of information to drive and speed up the process of discovering and adapting to the person's habits, preferences and needs. This article discusses the requirements posed by cultural competence on the knowledge management system of a robot. We propose a framework for cultural knowledge representation that relies on (i) a three-layer ontology for storing concepts of relevance, culture-specific information and statistics, person-specific information and preferences; (ii) an algorithm for the acquisition of person-specific knowledge, which uses culture-specific knowledge to drive the search; (iii) a Bayesian Network for speeding up the adaptation to the person by propagating the effects of acquiring one specific information onto interconnected concepts. We have conducted a preliminary evaluation of the framework involving 159 Italian and German volunteers and considering 122 among habits, attitudes and social norms.

## 1 Introduction

When Manuel Neuer was elected the best goalkeeper of the 2014 Football World Cup, and found himself face to face with his Chancellor Angela Merkel, he broke the protocol and spontaneously leaned over for a hug, that the Chancellor happily returned. Under very different circumstances, when Toyota's managing director Yuji Yokoyama had to announce a major recall campaign of their flagship cars due to braking problems, he bowed longer and deeper than usual to the officials of the Japanese Transport Ministry, to convey with the greeting his sincere apology for the whole situation.

The two episodes, reported in Fig. 1, well represent the depth and the ways of the influence of *culture* on a person's actions. Both Manuel Neuer and Yuji Yokoyama knew that meeting a government representative calls for a formal greeting (a handshake in Germany, a bow in Japan), assessed the context (a stadium moments after winning the Football World Cup, a press conference for the announcement of the recall campaign) and their own emotions (joy, shame) and modified the expected greeting gesture in a way that made their intentions immediately clear, evoking in the recipient the response they were hoping for.

The fact that Manuel Neuer and Yuji Yokoyama shared the same cultural background of their respective counterparts played a key role towards the success of their actions. Sticking out the tongue, as done by the elderly man of Fig. 2a, is considered rude and disrespectful in the USA, while in Tibet it is a formal greeting. Similarly, the two men shown head-to-head in Fig. 2b and who might appear as arguing to an Italian observer, are actually performing the *hongi*, the traditional greeting with which Māori people welcome a foreigner into their group.

Of course, culture is more than this. Besides greetings and facial expressions, culture influences individuals' lifestyles,

---

Barbara Bruno and Carmine Tommaso Recchiuto contributed equally.

✉ Barbara Bruno
   barbara.bruno@unige.it

   Antonio Sgorbissa
   antonio.sgorbissa@unige.it

1  University of Genoa, via Opera Pia 13, 16145 Genoa, Italy

2  Middlesex University Higher Education Corporation, The Burroughs, Hendon, London NW4 4BT, UK

3  Örebro University, Fakultetsgatan 1, S-70182 Örebro, Sweden

**Fig. 1** **a** Manuel Neuer hugging the German Chancellor Angela Merkel after winning the 2014 Football World Cup. **b** Toyota's managing director Yuji Yokoyama bowing to officials of the Japanese Transport Ministry before the press announcement of 2010 Toyota's recall campaign
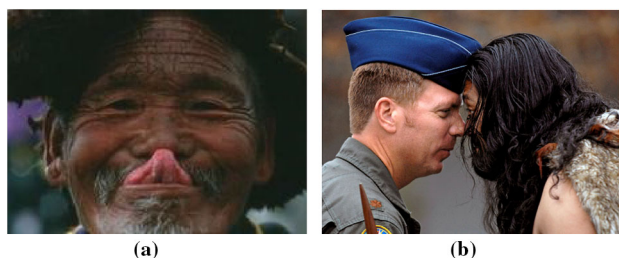


**Fig. 2** **a** Tibetan greeting. **b** Māori greeting

personal identity and their relationship with others both within and outside their culture. Culture is the shared way of life of a group of people that includes beliefs, values, ideas, language, communication, norms and visibly expressed forms such as customs, art, music, clothing, food, and etiquette. Cultures are dynamic and ever changing as individuals are influenced by, and influence, their culture by different degrees [26].

Various studies prove that culture also affects our interactions with, and expectations of, robots [12]. A survey about the requirements for a personal robot assistant conducted in 2005 reveals that people expect the robot to pay attention to what they are doing (85% of respondents), be polite (70%) and communicate in a human-like manner (71%) [9]. Interestingly enough, it is very difficult to meet those expectations without providing the robot with a certain level of *cultural competence*: "paying attention", i.e., understanding the meaning of gestures and words and reacting appropriately, is not possible without an understanding of the cultural identity of the person, as the episodes of Fig. 1 prove; similarly, as Fig. 2a shows, the definition of politeness is culture-dependent. Studies specifically focusing on the influence of the cultural background on the interaction with a robot reveal that people from different cultures not only have different preferences concerning how the robot should be and behave [12,24], but also tend to prefer robots better complying with the social norms of their own culture, both in the verbal [2,34] and non-verbal behaviour [11,18]. Such dif-

ferences affect the robot's likeability, as well as the trust, comfort and compliance it inspires [34].

While this problem is relevant in all applications requiring human-robot interaction, it is particularly critical whenever the robot is expected to be a companion for elderly [29], disabled people or children [8], who might, at once, need more and better assistance and be less capable of describing their needs and preferences. In such cases, the cultural competence of the robot has a tremendous impact on the quality of the care intervention, and even on its ethics [13].

Two complementary approaches have been proposed in the literature to tackle the problem of ensuring the cultural competence of a personal robot.

The "bottom-up" approach aims at adapting the robot's behaviour to suit the preferences and expectations of its user, under the assumption that any behaviour deemed as appropriate by a person is also appropriate for that person's culture. Examples of this interpretation range from a method for parametrizing the interpersonal distance and direction of approach on personal preference [31] to a complex framework for the learning and selection of culturally appropriate greeting gestures and words [32]. While this approach bypasses the problem of finding a suitable representation for the influence of culture on the robot's actions and perceptions, it is not well suited for encoding information expressed at national level, nor how such information might drive personal preferences.

On the contrary, the "top-down" approach relies on cultural information valid at national level (e.g., Hofstede's dimensions for the cultural categorization of countries [16]), to provide an informed *a priori* personal adaptation. Examples of solutions following the "top-down" approach include a system for the customization of the gestures and facial expressions of a virtual agent [28], and a framework for expressing the influence of culture on the gestures and words that a robot should use at a first meeting with a person [20]. The latter is among the very first attempts at merging the "top-down" and "bottom-up" approaches, by making use of empirical data (tagged video recordings) to complement the information given by Hofstede's dimensions. As the reported examples testify, the greatest limitation of the "top-down" approach is the difficulty in modelling the mapping between cultural information at national-level and variables defining the robot's behaviour, beside narrow, well-defined areas such as the interpersonal distance [6].

Moreover, both approaches seem to leave some areas uncovered: it is unclear, for example, how the "top-down" approach would allow for modelling and initializing symbolic variables (e.g., what are the eating habits of the user for breakfast? which holidays does she celebrate?) and rules (e.g., what is the appropriate behaviour for an invitation to dinner at a friend's house? what is the user's attitude towards

healthcare?), and it is certainly time-consuming, at least, to learn all such information with the "bottom-up" approach.

To address the problem of endowing personal robots with cultural competence on a broad spectrum of behaviours, we propose to draw inspiration from the field of *Transcultural Nursing* [26], which explores the influence of culture on the efficacy of care and proposes and validates culturally competent practices for (human) caregivers. We argue that, among the various angles from which the problem of defining culture and its influence on humans' behaviours is tackled, the "practical" perspective pursued in Transcultural Nursing is ideal for: (i) mapping human-related cultural knowledge onto robots' sensorimotor and verbal behaviours; (ii) defining metrics for the evaluation of the cultural competence of the resulting robot's behaviours and (iii) assessing the effect of the culturally competent robot onto the assisted person, in crucial aspects such as acceptability and efficacy.[1]

The contribution of this article is a hybrid "top-down/ bottom-up" software framework for the representation of heterogeneous cultural and contextual information required by a robot for elderly care to exhibit culturally competent behaviours. The framework relies on three core elements: (i) a three-layer ontology for storing all concepts of relevance, national-level information and statistics, person-specific information and preferences; (ii) an algorithm for the acquisition of person-specific knowledge, which uses national-level, culture-specific knowledge to drive the search, and (iii) a Bayesian Network for propagating the effects of acquiring one person-specific information onto interconnected concepts. To the best of our knowledge, this is the first framework for modelling the influence of culture on robot behaviours that can manage numerical and symbolic information and their combination, as well as rules and goals.

The hypothesis driving the first experimental evaluation of the proposed framework is that the amount of interactions required to learn individual preferences is significantly smaller when having national-level cultural knowledge about a user than in absence of such knowledge. Concretely, as detailed in Sect. 4, we assess whether, given a user who declares herself as belonging to one cultural group at national level (e.g., Italian), using the proposed framework and algorithms speeds up the acquisition of person-specific knowledge. The preliminary evaluation involved a total of 159 Italian and German volunteers. Planned future experiments will be devoted to assessing the perceived cultural competence of the robot, by adapting validated tools adopted in the field of Transcultural Nursing [21].

---

[1] This rationale is at the core of the H2020 project CARESSES (http://caressesrobot.org/), which aims at the development of culturally competent robots for elderly care. One of the key research areas of CARESSES is denoted as *Transcultural Robotic Nursing* [5], which is, ideally, the bridge between culturally competent human caregivers and culturally competent robot caregivers.
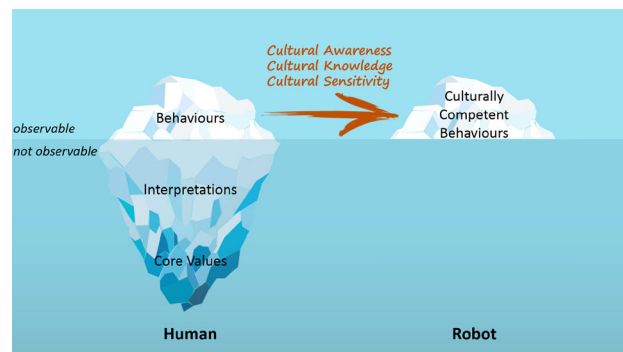


**Fig. 3** The *cultural iceberg model* (left) describes the relationship between a person's culture and behaviours, while the Papadopoulos, Tilki and Taylor model (center) describes the process allowing health practitioners to act with cultural competence. A framework for managing cultural knowledge is necessary for a culturally competent robot (right) to assess the actions and words of a person and respond accordingly

The article is organised as follows. Section 2 introduces the concept of culturally competent robot and details the requirements that cultural competence poses on the robot's knowledge management system. Section 3 describes the method we propose for meeting such requirements. Section 4 reports its implementation and experimental evaluation. Conclusions follow.

## 2 Motivations and Problem Statement

Figure 3 shows on the left-hand side the *cultural iceberg model*, which describes the relationship between a person's culture and behaviours, acknowledging the influence of the former on the latter. According to the model, inspired by the theories of the anthropologist Edward T. Hall [15], a person's cultural identity is composed of *core values* (at the bottom of the iceberg), their grounding in situations and events of everyday life (*interpretations*) and the *behaviours* that map the interpretations onto a person's physical and verbal capabilities. While the behaviours are immediately evident to an observer, the associated interpretations, as well as the underlying core values, are not directly observable and can only be inferred by correlating behaviours with generic knowledge and previous experiences.

The Papadopoulos, Tilki and Taylor model [25] has been devised by experts in Transcultural Nursing for developing culturally competent health (human) practitioners. The model consists of four constructs: *Cultural Awareness*, *Cultural Knowledge*, and *Cultural Sensitivity*, that lead to *Cultural Competence*. Let us again consider the Tibetan man of Fig. 2a: a culturally competent U.S. health practitioner, for example, would (i) understand that her interpretation of the gesture is influenced by her own culture and be aware that

the same gesture might have a different interpretation for the Tibetan man (cultural awareness), (ii) know that the gesture is a traditional formal greeting in Tibet (cultural knowledge) and (iii) respond ensuring that her actions and words enforce and convey trust, respect and empathy (cultural sensitivity). As a consequence, she might react by first mimicking the same gesture and then inquiring about its significance for the man.

As shown in Fig. 3 on the right-hand side, a robot does not have own core values and interpretations: its cultural awareness is therefore exclusively devoted to understanding the meaning, as intended by the person it is interacting with, of actions and words, as defined by experts in the field and adapted to a robot assistant. For the same reasons, also its cultural sensitivity is actually a-priori defined by experts in the field. Cultural knowledge provides the foundation to both stages, by storing the information required to understand the meaning of a person's actions and words, the information required to identify and perform an appropriate response, and finally the procedures to acquire new information and revising previous assumptions by directly interacting with the person. In our work, such information have been encoded by experts in Transcultural Nursing in the form of a corpus of Guidelines.[2]

The hybrid "top-down/bottom-up" software framework described in this article is specifically designed to manage such cultural knowledge.

Table 1 reports four interactions between a culturally competent assistive robot and, respectively, an Indian Hindu woman (first two blocks) and an English man (last two blocks), highlighting the role of cultural knowledge. The scripts, meant as a reference for development, have been written by experts in Transcultural Nursing and in accordance with the aforementioned Papadopoulos, Tilki and Taylor model for developing cultural competence [25]. From an implementation perspective, the cultural knowledge required by the robot can be divided into three categories.

Knowledge pertaining to the *context* includes information about the environment (e.g., allowing the robot to reach the *puja* table) and about the assisted person (e.g., allowing the robot to detect when Mrs. Chakrabarti is in a bad mood). In both cases, information can be static, a priori set (e.g., the location of the *puja* table), or dynamic, inferred from the robot's perception system (e.g., Mrs. Chakrabarti's mood).

Knowledge pertaining to the *robot*'s sensorimotor and communication capabilities is required by the robot to know what it can do and how the user might prefer it to be done. This knowledge again includes static, a priori information (e.g.,

describing the set of commands allowing the robot to perform the Namaste greeting, the associated parameters and their preferable values) and dynamic information (e.g., describing the robot's current posture and values of related parameters).

Knowledge pertaining to the *grounding of the core values* in the situation includes goals (e.g., leading the robot to chat with Mr. Miller about his past jobs, which triggers an open question, or to suggest to Mrs. Chakrabarti to walk with her to the *puja* table, which triggers the goal to reach another area of the house) and social norms (e.g., causing the robot to modify the values of the speech-related parameters to have a soft voice when it apologizes, or to reduce its speed when walking beside Mrs. Chakrabarti), which link the robot's behaviours to the context. Knowledge related to the grounding of core values which is not related the robot's actions is, in our proposal, straightforwardly mapped onto so-called *conversation subject matters*.

The above categories mention facts and preferences without distinguishing between person-specific knowledge and national-level knowledge, which is specific of a cultural group. However, as Table 1 shows, both are necessary to display a truly culturally competent behaviour. If the robot lacks person-specific knowledge, and thus relies on culture-specific, national-level knowledge only to tune its behaviour towards the assisted person, it is likely to end up having distorted, stereotyped representations of people (e.g., assuming that all British women have tea at five in the afternoon). Conversely, if the robot lacks culture-specific knowledge, it will either require a long and tedious setup phase, or incrementally add behaviours as they are learned, thus implying an unpredictably long phase in which it works with reduced functionalities. The situations in Table 1 provide a novel perspective on the problem: the robot tunes its behaviour either on person-specific knowledge (e.g., about Mrs. Chakrabarti, when it proposes to walk with her to the *puja* table, knowing that she has walking problems) or, *in the absence of it*, on culture-specific knowledge about the national culture (e.g., when it chooses to greet Mrs. Chakrabarti with the Namaste gesture since she is wearing a sari, or when it drives the discussion about Mr. Miller hobbies to the details of his past jobs, since the UK has a pragmatic mindset and under the assumption that Mr. Miller is at least familiar with it).

The use of culture-specific knowledge is key for the robot to make "educated guesses" about the likely appropriate course of action and ask confirmation about its intuitions, which, we hypothesise, speeds up the process of learning the preferences and customs of the assisted person without limiting the robot's capabilities, even at the earliest stages of deployment.

In short, the knowledge required by a culturally competent robot includes:

---

[2] The current corpus of Guidelines for Culturally Competent Robot Behaviours, together with a set of scenarios grounding them in daily life situations, is freely available at: http://caressesrobot.org/en/2018/03/08/caresses-scenarios-and-guidelines-available/.

**Table 1** Importance of cultural competence in daily-life situations: Mrs. Chakrabarti (India) and Mr. Miller (UK)

| Scenario | Robot skills | Cultural competence |
|---|---|---|
| MRS CHAKRABARTI: Hello Pepper! How are you? | | Culture-generic: PEPPER knows that greetings differ across cultures. |
| PEPPER: Oh, hello! *PEPPER looks at MRS CHAKRABARTI, recognises that she is wearing an Indian sari, pauses and greets her with Namaste* | Moving (Body, arms) Perception (Object / clothes recognition) | Culture-specific: PEPPER infers from her clothes that MRS CHAKRABARTI is Indian and chooses the greeting accordingly. |
| PEPPER: […] Could you please tell me your full name? | | |
| MRS CHAKRABARTI: My name is Purnima Chakrabarti. Can you say that? | | |
| PEPPER: Yes, of course I can, Purnima. *MRS CHAKRABARTI keeps quiet, she looks annoyed.* | Speaking (Catching key words) | Culture-specific: In India it is considered offensive to call older persons with their first names. |
| PEPPER: Oh, you look annoyed Auntie. *PEPPER does Namaste and slightly bows* | Perception (Emotion recognition), Moving (Body, arms) | |
| PEPPER: I made a mistake, I should not have called you like that. I am sorry. | | |
| MRS CHAKRABARTI: You know, Pepper, usually at this time of the day, I pray for my family and friends. If you excuse me, please. | | |
| PEPPER: Yes, of course. *(PEPPER moves slightly back as MRS CHAKRABARTI gets up from her chair.)* | Speaking (catching key words), Moving (body) | |
| PEPPER: Is it ok if I accompany you? | Speaking (asking yes/no questions) | Person-specific: PEPPER knows that MRS CHAKRABARTI has difficulties with walking and offers to help. |
| MRS CHAKRABARTI: Yes, I would like that. | | |
| PEPPER: I can walk next to you and you can put your hand on my shoulder if you would like. | | Culture-specific: PEPPER knows that puja is a prayer ritual performed by Hindus, usually in a dedicated area of the house. |
| MRS CHAKRABARTI: Thank you, Pepper. *(PEPPER leads MRS CHAKRABARTI to the puja table.)* | Perception (objects localization), Moving (indoor navigation) | Person-specific: PEPPER knows where the puja table is located in this house. |
| MR MILLER: I like machines and I like to fix things, so I am excited to have you. | | |
| PEPPER: Oh, that is very nice. I would like to learn more about you and your hobbies. | Speaking (Catching key words) | |
| MR MILLER: Well, let's see… I worked in manufacturing all my life, in different industries. First in the railways, then telecommunications and before retirement in solar! | | Culture-specific: The UK has a pragmatic orientation. |
| PEPPER: Oh, sounds exciting! Which one was the most interesting job? | Speaking (Asking open questions) | |
| MR MILLER: I am a cricket fan and England is playing today at The Oval with Sri Lanka. I do not want to miss it. | Speaking (Catching key words) | Culture-specific: Cricket is a popular sport in the UK. |
| PEPPER: Of course, let me check what time the game starts. *(PEPPER shows the TV time table)* | Connecting to the internet | |
| MR MILLER: Oh, that is lovely. Can you come closer, so I can read? *(PEPPER moves closer)* | Moving (Approaching person) | Person-specific: PEPPER knows the distance and volume preferred by MR MILLER. |

– culture-generic knowledge about the context, the robot itself and the grounding of core values, i.e., knowledge ideally comprising all concepts from all cultures with no information on how the former relate to the latter;
– culture-specific, national-level knowledge, describing the cultural background of the assisted person, that the robot can rely on whenever specific information is not available;
– person-specific knowledge, describing the way in which the cultural identity, preferences and environment of the assisted person shape the appropriate robot behaviours.

The knowledge must be complemented by methods for the smart integration of person-specific and culture-specific knowledge, which rely on the latter to drive the discovery of the former.

## 3 Proposed Method

An *ontology* is a formal naming and definition of the types, properties, and interrelationships of the entities relevant for a particular domain of discourse [14]. The terminology defining the domain of discourse, containing general properties of concepts, is stored in the terminological box (TBox) of the ontology, while knowledge that is specific to instances belonging to the domain is stored in the assertional box (ABox) of the ontology. Ontologies allow non-technical users to easily[3] encode knowledge about the domain, which is a key property in cross-disciplinary contexts, such as ours.

We represent the knowledge required by a culturally competent robot with a modular ontology structure composed of an upper ontology and a number of domain-specific ontologies.

*Upper ontologies* have been proposed to support semantic interoperability among different domain-specific ontologies, and consist of very general terms that are common across all the considered domains, thus providing a common starting point for the formulation of definitions [23]. Terms in domain ontologies are ranked under the terms in the upper ontology. At the same time, a number of *domain-specific ontologies* have been already developed to describe highly-specific domains that are likely to be connected with many others, such as the Time ontology[4] proposed as a standard for the Semantic Web by the W3C, or the Food, Politics, Sport and Wildlife ontologies included in the BBC collection.[5]

We adopt the OWL-2 language [33] to describe the ontology. In the OWL-2 formalism, the TBox is composed of
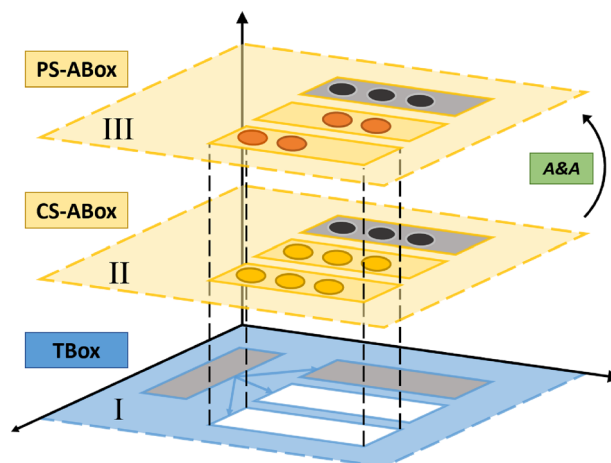
**Fig. 4** Knowledge representation architecture for a culturally competent robot. The TBox layer (I) includes terms from existing upper and domain-specific ontologies (grey boxes) and ontologies modelling cultural-knowledge that we propose (white boxes). The Culture-Specific ABox layer (II) includes instances (yellow circles) encoding knowledge at national-level, while the Person-Specific ABox layer (III) includes instances (orange circles) encoding knowledge uniquely related to the user. Some instances of existing ontologies (dark circles) may not change between the two ABox layers

*classes* and *properties*, which include *data properties*, relating instances of a class to literal data (e.g., strings, numbers), and *object properties*, relating instances of a class to other instances. *Instances* of classes and properties are stored in the ABox.[6]

The relationship between the TBox and the ABox of the ontology is sketched in Fig. 4. The Figure describes four core elements:

– *Culture-generic knowledge*, a layer that stores the terminology (TBox - I) required to represent all the information related to the context, the robot, and the grounding of the core values, ideally *for all* the cultures of the world;
– *Culture-specific settings*, a layer that stores the assertions (CS-ABox - II) required to represent cultural information at national level;
– *Person-specific settings*, a layer that stores the assertions (PS-ABox - III) required to represent the unique cultural identity, preferences and environment of the assisted person;
– *Assessment & Adaptation*, an algorithm (A&A) for the discovery of person-specific settings in light of culture-specific settings, e.g., relying on "educated guesses" to be confirmed through dialogue or autonomous robot observation.
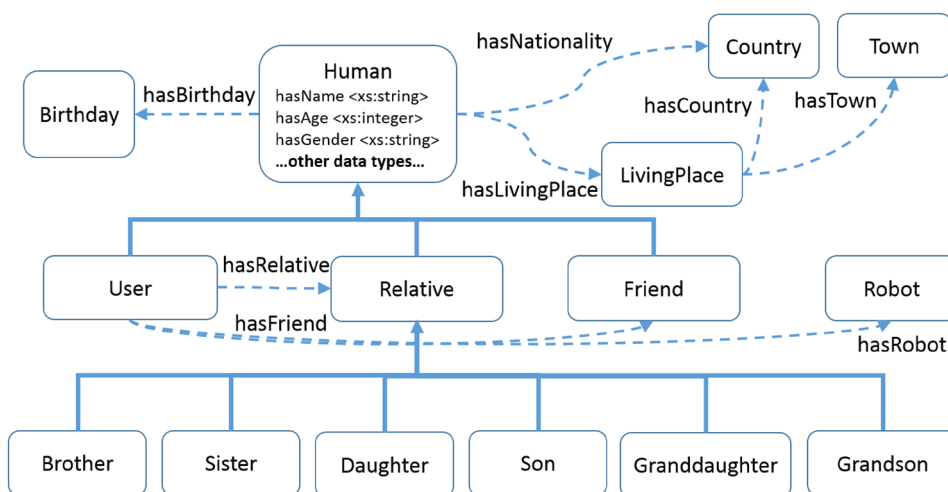
**Fig. 5** User TBox
(partial)—Only some classes
and properties are shown



It is easy to see that there may be concepts (e.g., the definition of "woman", or "day") for which we do not need to create different instances in the culture-specific and person-specific layers. Such instances are ignored by the Assessment & Adaptation algorithm.

### 3.1 Culture-Generic Knowledge

As discussed in Sect. 2, the knowledge required by our application includes:

1. context-related information, describing (1) the assisted person and (2) the environment;
2. robot-related information, describing (3) the actions that the robot can perform, (4) their parameters and, eventually, (5) their combination into higher level planning operators;
3. information related to the grounding of core values, describing (6) goals, (7) social norms and (8) conversation subject matters.

#### 3.1.1 Context Domain

Figure 5 shows a portion of the TBox defining the User domain. In the Figure, boxes denote classes (e.g., User, Human), solid lines denote hierarchical "is a" relationships (e.g., User is a Human), and dashed lines denote object properties (e.g., hasRelative). Data properties (e.g., hasAge) appear within the box of the class they refer to. Since User is a Human it inherits from its parent class the data properties hasName, hasAge, hasGender, as well as the object properties hasBirthday, hasNationality, hasLivingPlace. In addition, it may be related to other classes through specific object properties such as hasRelative and hasFriend, which define the network of people whom the robot is expected to meet or know about and their relation with the User. Among

the other properties, User is characterized by having a Robot (or more): in a simplified Description Logics formalism [3] this is expressed as:

$$\text{User} \sqsubseteq \text{Human} \sqcap \exists \text{hasRobot.Robot} \qquad (1)$$

Figure 6 shows a portion of the TBox defining the Environment domain, i.e., a person's house with those furniture, appliances and objects within it, which are of relevance for the interactions (e.g., because they are strong indicators of a person's cultural identity, such as the Tatami, or tightly connected to habits and preferences, such as the TeaCupSet and the Coffeemaker). This knowledge also serves as a reference for the robot's perception system, allowing for linking static, semantic information to dynamic, numerical data (not described in this article). It is important that the descriptions of the Environment and all other domains are not limited to relevant or common classes and properties for one nation and culture, but rather include concepts from many countries and cultures. Although it is surely unlikely that an elderly English woman sleeps on a tatami, it is not impossible: cultural competence demands that the caregiver is able to accept such a possibility, and act appropriately.

#### 3.1.2 Robot Domain

The design of an ontology for the representation of robot tasks is a complex and open issue, tackled, for example, by a dedicated IEEE Working Group [17], and it goes beyond the scope and goals of this article. In the present work, we exclusively focus on the representation of those elements related to the robot's behaviour which depend on cultural factors (shown in Fig. 7).

Let us consider the action ApproachUserAction, that describes an atomic sensorimotor behaviour that the robot shall perform to move from one location to another, close to

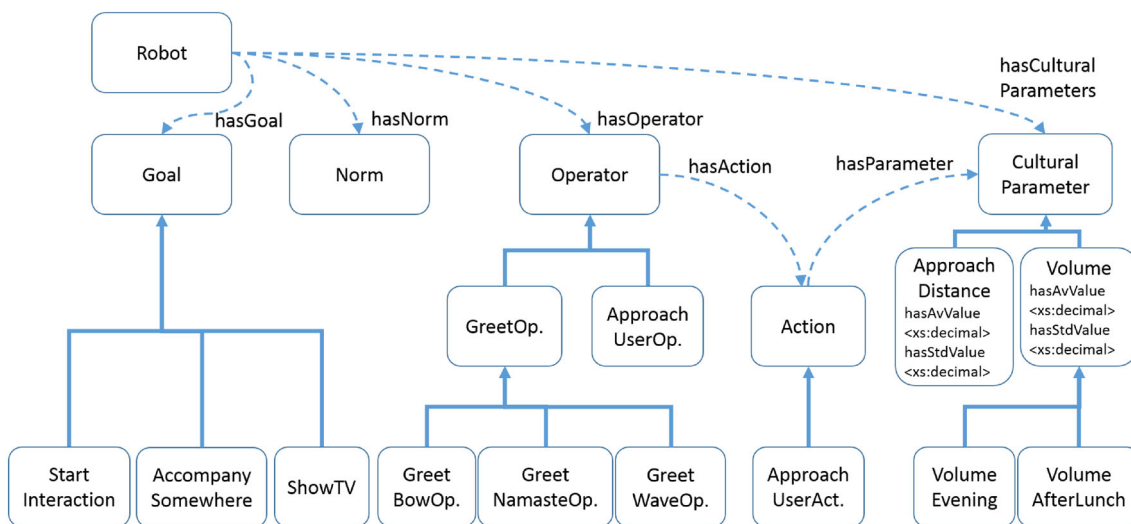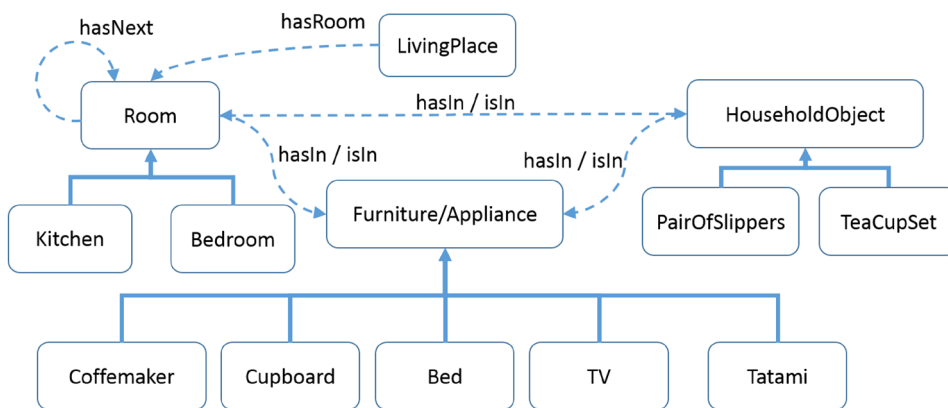**Fig. 6** Environment TBox (partial)—Only some classes and properties are shown



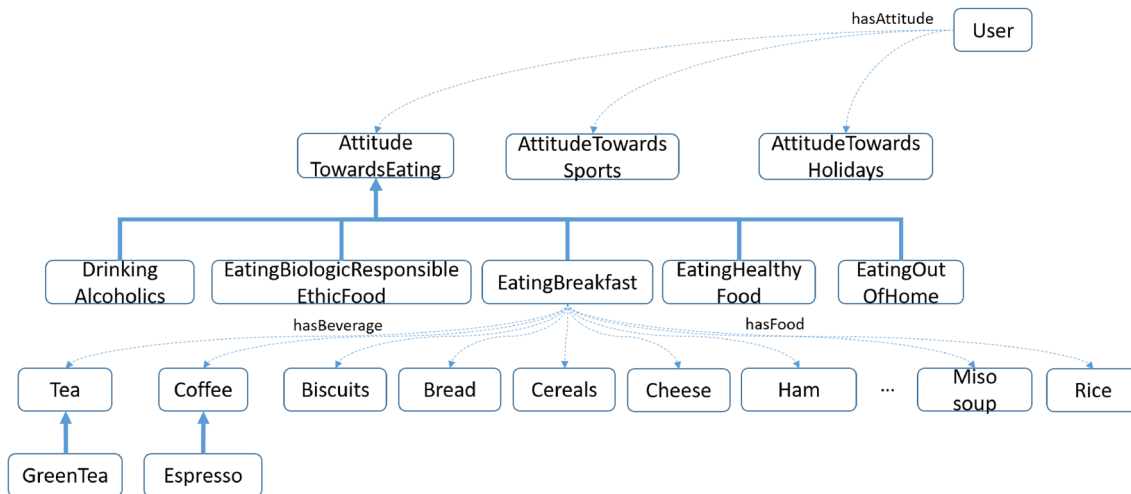**Fig. 7** Robot TBox (partial)—Only some classes and properties are shown



**Fig. 8** Conversation Subject Matters TBox (partial)—Only some classes and properties are shown

where the user is. The action has a number of parameters, including final location and final distance from the person. While the former is contingent and tightly related to the task at hand, it is easy to see that the latter might change in accordance with cultural and personal preferences [6]. We leave the representation of the culture-independent elements required for planning to suit the requirements of the chosen planner [19], and represent culture-dependent parameters of actions with the class CulturalParameter: the object property hasParameter relates actions to each cultural parameter they have. Each cultural parameter is represented as a subclass of CulturalParameter, which has a number of data properties to specify admissible values and semantic meanings associated with those values (e.g., allowing for defining a certain range of Volume values as "low"). The fact that the same parameter might have different preferred values in different situations (e.g., someone living in a condominium might want the robot to lower its Volume in the evening not to disturb the neighbours) is modelled by building a collection of subclasses below the parameter, with one class per situation of relevance (e.g., VolumeEvening). The taxonomy of subclasses corresponding to different situations and their initial values are defined by experts[2] and revised through interaction [27].

State-of-the-art planners [19] typically group actions into higher-level planning operators, which represent more complex robot behaviours. As an example, let us consider the GreetOperator, which requires the robot to perform a greeting gesture and utter an appropriate sentence. The gesture, the sentence, and the relation between the two are all culture-dependent: we represent each planning operator as a subclass of Operator, which is linked to actions via the object property hasAction. Variants of an operator are modelled as a collection of subclasses of the operator (e.g., GreetBowOperator, GreetNamasteOperator and GreetWaveOperator represent three different greetings adopted in different cultures across the world). The mechanism we adopt to let the planner know which operator is to be preferred with a specific culture, or person, is described in Sect. 3.2.

### 3.1.3 Core Values Domain

Goals, i.e., objectives driving the robot's behaviour, are modelled as subclasses of the class Goal and expressed in the planner formalism as a desired state that the robot should achieve. As for actions and planning operators, culture-independent properties of goals and norms are not shown in the Figure. As an example, seeing the assisted person entering the room where the robot is might trigger the goal StartInteraction, which requires the robot to offer its assistance to the user, while a specific request from the user might trigger the goal ShowTV (as it happens with Mr Miller, see Table 1). All the goals that the robot shall be able to accept must be described in the ontology. During interac-
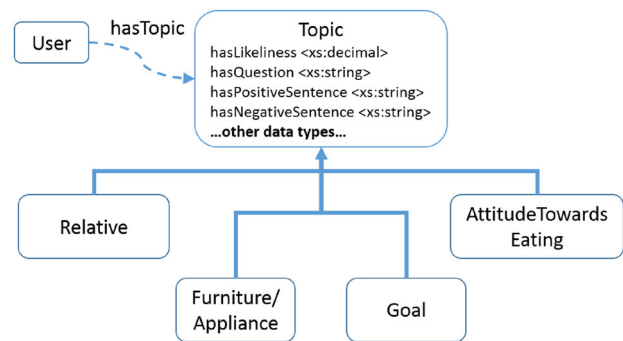


**Fig. 9** The classes Topic and User and the property hasTopic allow for storing culture-specific and person-specific information and relate it to other concepts

tion, the robot uses this information to trigger or suggest goals to be achieved depending on direct requests or the cultural knowledge it has about the person (e.g., the goal AccompanySomewhere is proposed as the robot detects that Mrs Chakrabarti is heading to the puja room for prayer, see Table 1).

Social norms represent additional constraints relating goals, planning operators, actions and cultural parameters with specific contexts. Concretely, they define additional goals that must be met, specific situations (states) that must be achieved/avoided, planning operators, actions or values of the cultural parameters which must be chosen/avoided in a specific situation. Norms are expressed in the planner formalism and modelled in our ontology with the class Norm. As for planning operators, the mechanism we adopt to let the planner know which goals and norms are suitable for a specific culture, or person, is described in Sect. 3.2.

Figure 8 shows a portion of the TBox defining the Conversation Subject Matters domain, intended as the collection of knowledge which is meant at keeping the interest of the user and show the robot's attentiveness to the person's values, preferences, beliefs, etc. Figure 8 focuses on the terms describing the user's AttitudeTowardsEating, AttitudeTowardsSports and AttitudeTowardsHolidays, which are the ones considered during the experimental evaluation. Specific habits and preferences are modelled with subclasses, such as EatingBreakfast, with object properties such as hasBeverage and hasFood relating the preference/habit to actual objects (e.g., drinks and food). As already stated, the TBox should represent concepts (e.g., drinks and food) that are typical of as many cultures as possible, whichever the nationality of the user, to avoid stereotypes. Luckily, many of such concepts (e.g., all possible beverages) are part of existing domain ontologies that are imported in our representation.

While some preferences and attitudes can be related to goals and social norms (e.g., a conversation about eating habits occurring in the late afternoon leads the robot to ask

the person whether she wants assistance for preparing dinner), most of them are only used for "chit-chatting", under the intuition that users might appreciate a robot that is familiar with the very same concepts they are familiar with.

The above principle forces the robot to tune in on the user's preferences *concerning the robot*, which does not necessarily mean that it will end up mimicking the assisted person. Concretely, the fact that a user is Italian, for example, does not constrain the robot to behave as an Italian, or as expected with an Italian; the robot will rather act in accordance with its knowledge of culture-specific Italian habits *at the beginning*, and progressively change its behaviour as it discovers how the person likes it to be.

### 3.2 Culture-Specific Settings

Figure 9 shows the solution we propose to store information about how all aforementioned concepts are related to culture-specific (national-level) and person-specific (user) preferences and settings.

The class User represents the person assisted by the robot, that is related to all the concepts described in Sect. 3.1 by ownership (e.g., of objects and furniture), preferences, habits, beliefs, etc. Instances of User can be of two types: culture-specific instances (CS-ABox layer in Fig. 4) are used to store information about national-level culture, while person-specific instances (PS-ABox layer in Fig. 4) describe real people assisted by the robot. The class Topic is a super-class to all classes in the context, robot and grounding of core values domains (see Figs. 5, 6, 7, 8). Its data property hasQuestion contains the question(s) the robot should use to ask the user about any instance subsumed by Topic (e.g., "Is it ok if I stand this close to you?" for *English-speaking* instances of the class ApproachDistance), while the data properties hasPositiveSentence and hasNegativeSentence contain sentences that the robot can use to express, respectively, a positive or a negative attitude towards the instance subsumed by Topic (e.g., hasPositiveSentence for an instance of the class Kitchen might be "The kitchen is the heart of a home!", while hasNegativeSentence for an instance of the class AttitudeTowardsSports, borrowed from the actress Phyllis Diller, could be "My idea of exercise is a good brisk sit!"). All sentences, and especially negative ones, should be checked by experts, to ensure that they are ethically and culturally sound.

**Definition 1** The likeliness[7] $l(a)$ of an instance assertion $a$ is a value in the range [0, 1], associated with the assertion $a$. It

corresponds to a reasonable estimate, to the best of available knowledge, of the a posteriori probability of the assertion $a$.

In the culture-specific ABox layer describing culture $C$, the data property hasLikeliness is filled with the probability $l(a)$ that assertion $a$ (an instance of Topic) holds for a person, given that we know that she belongs to culture $C$. To clarify the concept, let us assume that the chances that a British person does some sport are quite high. This information might be represented in the culture-specific ABox as:

$$
\begin{aligned}
&\text{User(GB\_GEN)}\\
&\text{DoesSport(GB\_DOES\_SPORT)}\\
&\text{hasTopic(GB\_GEN, GB\_DOES\_SPORT)}\\
&\text{hasLikeliness(GB\_DOES\_SPORT, 0.7)}
\end{aligned}
\tag{2}
$$

which corresponds to saying that there exists a culture-specific instance GB_GEN of the class User (representing British culture at national level) and an instance GB_DOES_SPORT of DoesSport ⊑ AttitudeTowardsSports, that the second is a filler of the former for the property hasTopic (which allows the robot to use the sentences in GB_DOES_SPORT, presumably in English), and that the data property hasLikeliness of GB_DOES_SPORT is set to 0.7.

**Definition 2** In the culture-specific layer, we define the likeliness $l(a)$ as depending only on the national culture $C$ and hence ideally corresponding to the conditional probability of assertion $a$ given the evidence of $C$ : $p(a|C)$.

Beside the mathematical definition, the likeliness has a practical meaning which might change for different classes. With no loss in generality, we can define a hierarchy of object properties subsumed by hasTopic, highlighting the different meanings in which the user/instance relation is intended. For example, User might be related to instances of classes in the Environment domain by the property hasOwnership, to instances of classes in the Robot domain by hasPreference, to instances of classes in the Conversation Subject Matters domain by hasHabit, hasBelief, hasAttitude, which are all derived from hasTopic.

The use of a comprehensive culture-generic TBox and a culture-specific ABox describing the relation between a given culture and all the elements defined in the TBox allows for avoiding *stereotyped representations* of cultures. It is a well know fact that "biscotti (cookies) are commonly eaten for breakfast in Italy,"[8] but, although this is probably true for many Italian men and women, it is not valid for all of them. While the stereotype simply assumes that what is valid for

---

[7] We introduce the term *likeliness* for two reasons: (i) to highlight the fact that it is not necessarily the result of statistical analyses, but it can also be provided by experts on the basis of qualitative assessment; (ii) to provide a unique name for the a posteriori probability (see Definition 1), the conditional probability (see Definition 2) and the evidence (see

Definition 3), which our algorithms for the Assessment & Adaptation (see Sect. 3.4) use concurrently.

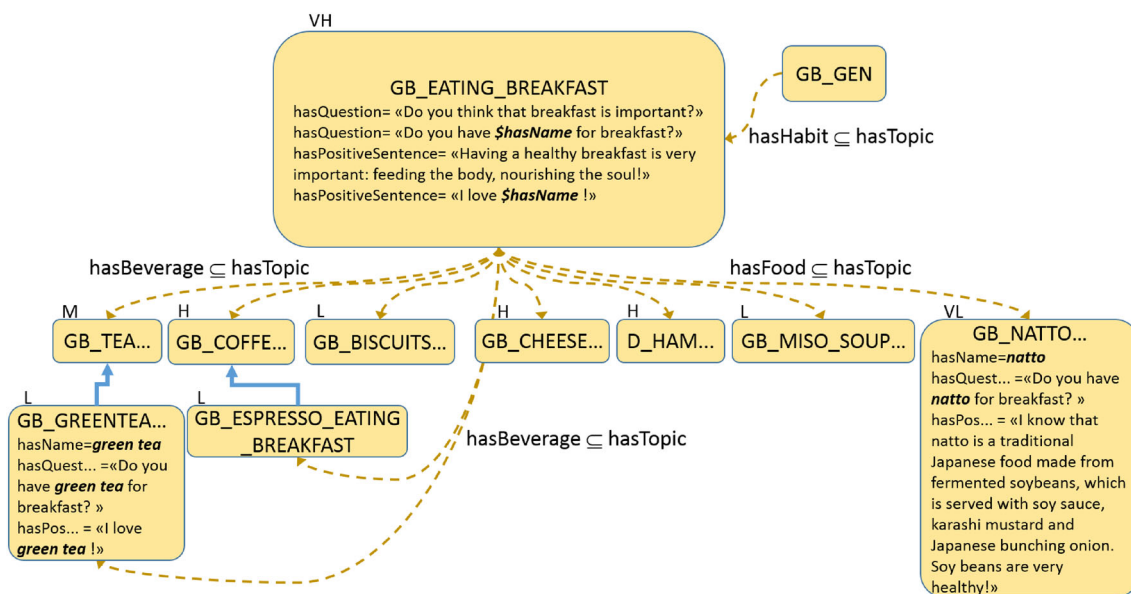[8] https://en.wikipedia.org/wiki/Breakfast#Italy

**Fig. 10** ABox describing British culture-specific (GB prefix) breakfast habits

most is valid for all, our culture-specific layer specifies the likeliness of many different food to be eaten for breakfast by an Italian person. This means that not only the culture-specific layer is truly representative of all the facets of a culture, but also that it allows individuals belonging to a culture to stray away from its *most likely* options as far and as many times as they want.

Figure 10 shows the portion of the culture-specific ABox of GB_GEN related to breakfast habits and preferences.[9] In the Figure, boxes denote instances of classes (e.g., GB_GEN, GB_EATING_BREAKFAST), yellow dashed lines denote assertions of object properties (e.g., GB_BISCUITS_EATING_BREAKFAST is a filler of GB_EATING _BREAKFAST for the property hasFood). Data properties (e.g., hasQuestion) appear within the box of the instance they refer to, while hasLikeliness values appear on the top-left corner of the instance they refer to and are denoted with literals instead of numbers, with 0.05 mapped to Very Low (VL), 0.1 to Low (L), 0.2 to Medium (M), 0.4 to High (H), 0.7 to Very High (VH). The reason for this choice is practical: while it is very difficult to obtain precise likeliness values from statistical analyses, it is much easier to infer approximate, qualitative values from the vast (but often inhomogeneous) corpus of information in the literature and on the web (see Sect. 4). A discrete representation of likeliness values makes it easier to merge approximate and precise values. Lastly, blue solid lines are used to remind the reader of existing hierarchical relationships between the classes that the instances belong to (e.g. in the TBox, GreenTea *is a* Tea).

Figure 11 shows the portion of the culture-specific ABox of GB_GEN related to robot goals, actions and cultural parameters. Likeliness values are used to specify how appropriate each instance is for the British culture, and guide the decisions of the planner which ultimately determines the robot's behaviour. As an example, if the situation calls for a greeting, the robot will execute the operator GB_GREET_WAVE_OPERATOR, since it has a higher likeliness than all other available greeting operators. Similarly, whenever executing the action GB_APPROACH_USER_ACTION it will set its parameter *approach distance* to the range of values specified by GB_LONG_APPROACH_DISTANCE, which is the most likely setting among available ones. Lastly, the goal GB_WEATHER_FORECAST, having high likeliness for the British culture, is likely to be pro-actively suggested by the robot as a service it can provide.

Instances of classes are created so that each instance is filler for *no more than one* object property derived from hasTopic, and its name is guaranteed to be unique by including the name of the instance itself and the one whose property is filled (as in GB_BISCUITS_EATING_BREAKFAST). This means that, by considering all instances of Topic and property assertions derived from hasTopic, the culture-specific ABox layer is a tree rooted in the corresponding instance of User (e.g., GB_GEN in Figs. 10, 11). This constraint is key for storing into instances unambiguous contextual information about their predecessors in the tree, e.g. to distinguish between "biscuits that the person may or may not eat for breakfast" (i.e., the instance GB_BISCUITS_EATING_BREAKFAST) and "biscuits that the person may or may not have with tea in the afternoon"

---

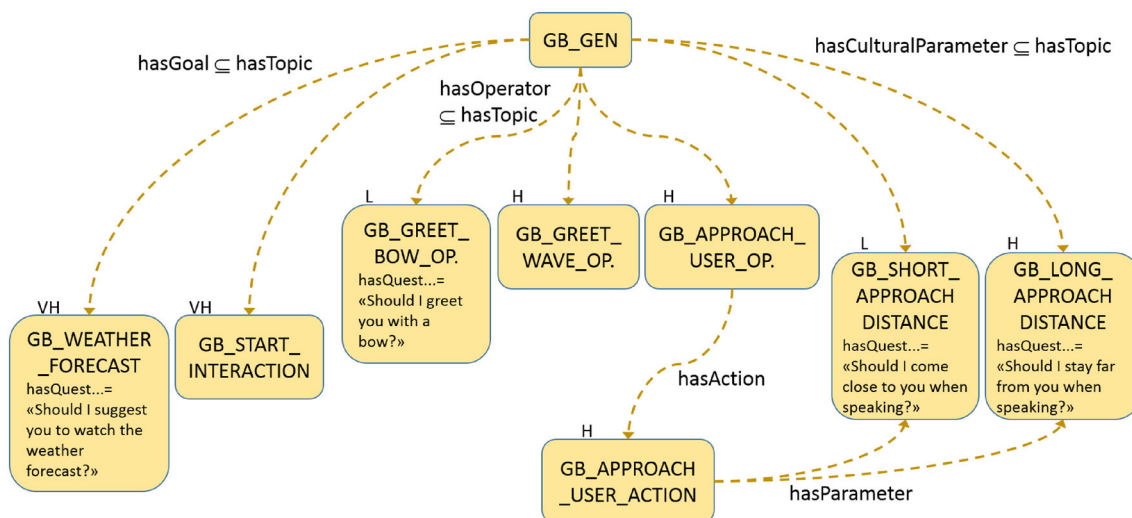[9] As usual, only object and data properties that are relevant for the discussion are shown.

**Fig. 11** ABox describing British culture-specific (GB prefix) robot goals, actions and cultural parameters

(i.e., the instance GB_BISCUITS_EATING_AFTERNOONTEA, not shown in the Figure), and even "biscuits that the person may or may not need to buy". This constraint is exploited by the Assessment & Adaptation algorithm (see Sect. 3.4) to ensure that the robot does not give wrong interpretations to the person's statements.

The sentences stored in the data properties hasQuestion, hasPositiveSentence, and hasNegativeSentence ensure that the robot can discuss the instance they refer to with the user.

We adopt two mechanisms to fill the data properties above. A number of complete sentences (such as "Having a healthy breakfast is very important: feeding the body, nourishing the soul!" in the instance GB_EATING_BREAKFAST) are encoded at setup time by the designer and validated by experts[2]. As videogame designers know, dramatization is key for improving the user's experience, and it can hardly be achieved through automatic composition of sentences [4,30]. However, manually encoding all verbal utterances is very time consuming. As a backup solution, we rely on simple automated composition mechanisms, which exploit the hierarchical structure of the ontology and the unique connections between instances defined by the property hasTopic. As an example, in Fig. 10 the instance GB_EATING_BREAKFAST encodes the hasQuestion "Do you have $hasName for breakfast?", which is automatically copied and filled in all the instances that are filler of GB_EATING_BREAKFAST along the property hasTopic (e.g., GB_GREENTEA_EATING_BREAKFAST and GB_NATTO_EATING_BREAKFAST), by using the value of the corresponding data property hasName.

## 3.3 Person-Specific Settings

The core element of the person-specific ABox layer is the instance of User which corresponds to the real person assisted by the robot, e.g., User(DOROTHY_SMITH). All instances of Topic and its subclasses connected to User(DOROTHY_SMITH) belong to the person-specific ABox layer and uniquely refer to that specific user.

**Definition 3** In the person-specific layer, the likeliness $l(a)$ corresponds to the evidence of assertion $a$ collected through interaction with the user.

Concretely, evidence about Mrs. Smith's habits concerning sports may be represented as:

$$
\begin{aligned}
&\text{User(DOROTHY\_SMITH)}\\
&\text{hasSpecific(GB\_GEN, DOROTHY\_SMITH)}\\
&\text{Does\_sports(DS\_DOES\_SPORT)} \qquad\qquad (3)\\
&\text{hasSpecific(GB\_DOES\_SPORT, DS\_DOES\_SPORT)}\\
&\text{hasLikeliness(DS\_DOES\_SPORT, 0)}
\end{aligned}
$$

Notice that, in the person-specific ABox layer, instances of Topic do not need to be directly linked to the user through property instances of hasTopic, as they are simply fillers of the corresponding instances in the culture-specific ABox layer for the hasSpecific property.

Figure 12 shows an example of an ABox including both the culture-specific layer (yellow boxes) and the person-specific layer (orange boxes). In the Figure, connections between instances in the two layers (through the hasSpecific property) are represented by overlapping the boxes (i.e., without a corresponding arrow).

Instances are inserted in the person-specific layer at two different times: in the *setup* phase, engineers, experts, the
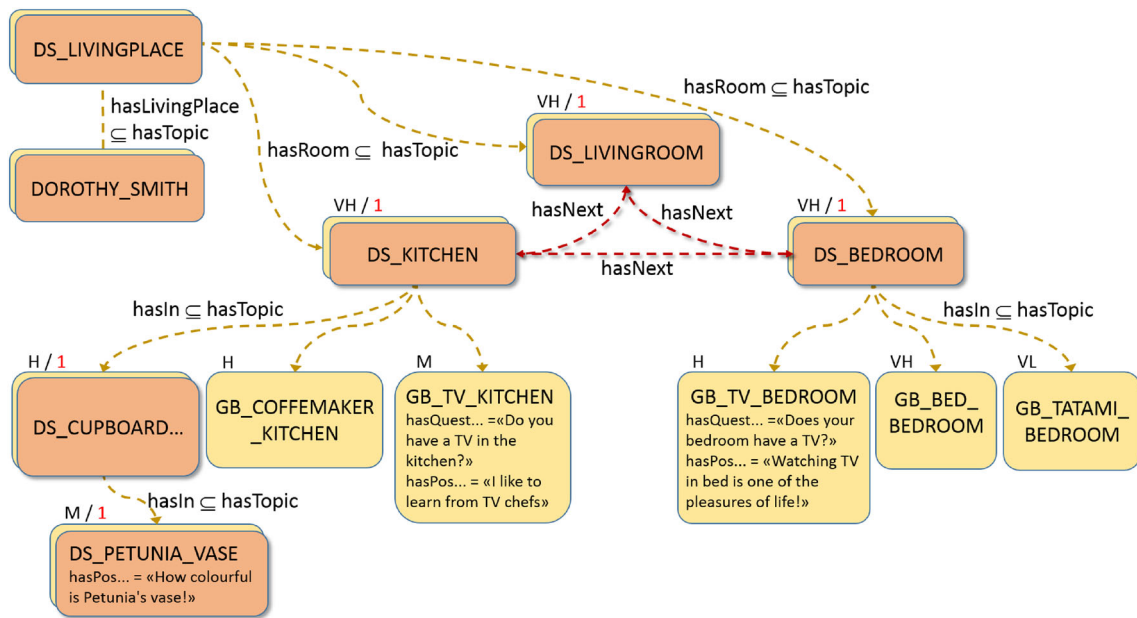
**Fig. 12** ABox describing British culture-specific (GB prefix) and person-specific (Dorothy Smith) knowledge

assisted person, caregivers and relatives, add knowledge that is a priori available about the user (e.g., which rooms are part of her house and how they are connected to each other), using ad-hoc prepared tools and tutorials aimed at facilitating the insertion and validation of knowledge; at *run-time*, the robot autonomously acquires knowledge either from its own perceptual system or from interactions with the user.

Notice that property instances of hasTopic are only present in the culture-specific ABox layer (yellow arrows), whereas other property instances are only present in the person-specific ABox layer (red arrows). For example, in Fig. 12 property instances of hasNext are only specified at the person-specific level (because they have little to no meaning at culture-specific level), to connect instances of Room with each other to represent the topology of the specific house of the user. As a consequence, the system can talk about rooms, but not about their topological relationships.

As Figure 12 shows, at the end of the setup phase the ontology lacks person-specific knowledge for many instances. In all such cases, the robot must assess at run-time person-specific knowledge (and adapt to it) by using generic knowledge as a starting point.

Finally, notice that the black likeliness values in Fig. 12 refer to the culture-specific layer, whereas the red values refer to the person-specific layer: English people might have a Medium probability of keeping a vase in the cupboard, but we know for sure that Dorothy Smith has one, since someone (Mrs. Smith herself, or a relative) added this piece of information during setup.
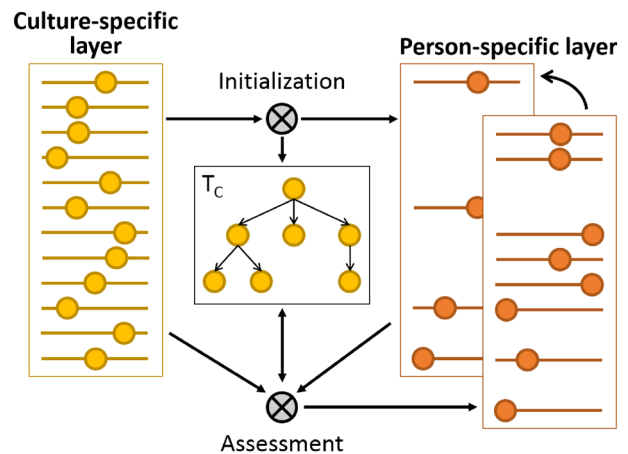


**Fig. 13** The offline *initialization* and online *likeliness-driven assessment* phases, which, respectively, create and update the person-specific ABox layer. Notice that only one culture-specific layer is used

## 3.4 Likeliness-Driven Assessment

The goal of the assessment and adaptation phase is to learn the person-specific likeliness values (i.e., evidence) for all instances of relevance for the interactions. Notice that in some cases the likeliness in the person-specific layer of the ABox will be 1 or 0, as the user, for example, either has a TV in the bedroom or not, while instances related to preferences or habits might lead to a more varied output. Moreover, different methods for the assessment of person-specific settings might have different reliability (e.g., directly asking the user guarantees a more reliable assessment than autonomously inferring information from sensor data) and such differences

**Algorithm 1** Initialization

---
**Require:** $A_C$
**Ensure:** $T_C$
1: $T_C \leftarrow \emptyset$
2: $L \leftarrow \emptyset$
3: $r \leftarrow \text{init}(T_C)$
4: $L \leftarrow \text{updateTopicsTree}(r, T_C, A_C)$
5: **while** $L \neq \emptyset$ **do**
6:     $L_{next} \leftarrow \emptyset$
7:     **for all** $a^* \in L$ **do**
8:        $L_{next} \leftarrow L_{next} \cup \text{updateTopicsTree}(a^*, T_C, A_C)$
9:     **end for**
10:    $L \leftarrow L_{next}$
11: **end while**

---

**Algorithm 2** Assessment

---
**Require:** $A_C, T_C, A_U$
**Ensure:** $A_U$
1: $a^* \leftarrow \text{findMax}(A_U, A_C, T_C)$
2: $f^* \leftarrow \text{getParent}(T_C, a^*)$
3: **while** $f^* \notin A_U$ **do**
4:     $a^* \leftarrow f^*$
5:     $f^* \leftarrow \text{getParent}(T_C, a^*)$
6: **end while**
7: $l(a^*) \leftarrow \text{assess}(a^*)$
8: createSpecificInstance($l(a^*), a^*, A_U$)
9: update($l(a^*), a^*, T_C$)

---

can be embedded in the person-specific likeliness values and the way they are handled.

In our work, we assume that the robot acquires person-specific knowledge by directly asking the user, using the data property hasQuestion associated to all instances of Topic. The simplest assessment procedure, trivially, is to go through the instances one by one, without using culture-specific information, which might lead the robot to ask Mrs. Smith whether she likes Miso soup (a Japanese dish) or Lasagne for breakfast.

We propose an algorithm, sketched in Fig. 13 which consists of an offline *initialization* phase and an online *likeliness-driven assessment* phase.

The *initialization* phase is composed of two steps. In the first step, the person-specific ABox $A_U$ is populated with the knowledge that is available at setup by experts, relatives, etc. At the end of this step, as shown in Fig. 12, some instances of the culture-specific ABox $A_C$ are exactly replicated in $A_U$, others have a corresponding instance in $A_U$, with different likeliness values and other data properties, and others do not have a corresponding instance in $A_U$. In the second step, the assessment tree $T_C$ is built from the culture-specific ABox $A_C$ to drive the discovery of missing person-specific knowledge.[10]

Algorithm 1 details the second step. The routine init($T_C$) initializes the *assessment tree* with the culture-specific instance of User as root and returns it in $r$ (line 3). Then all the instances directly connected to $r$ through a property derived from hasTopic are added to the tree by the function updateTopicsTree(...), one level below root, and are returned in the set $L$ (line 4). In subsequent iterations (lines 5–11), each instance $a^*$ in $L$ is considered, and all instances that are fillers of $a^*$ for a property derived from hasTopic are added to the tree, at the level below $L$ until $L$ is empty. If two

instances $l_1, l_2 \in L$ belong to classes $C_1$ and $C_2$ that have a child/parent relationship $C_2 \sqsubseteq C_1$ in the TBox, then $l_1$ is directly linked to $a^*$ in the tree, whereas $l_2$ is linked to $l_1$ and not to $a^*$. Concretely, updateTopicsTree(...) ensures that the instance that belongs to the superclass is the closest to the root in the tree.

At the end of Algorithm 1, there is a structure, the *assessment tree*, which stores the knowledge in the culture-specific layer in the form of a tree rooted in the culture-specific instance of User and uses the object property hasTopic, and the hierarchical relationships among instances, to define branches. As discussed in Sect. 3.2, the tree-like structure is key to ensure that each instance has a unique context. As an example, Fig. 14 shows the *assessment tree* corresponding to the culture-specific ABox of Fig. 10.

The online *likeliness-driven assessment* phase follows the steps sketched in Algorithm 2 to identify the instances to discuss with the user, and update the person-specific ABox accordingly.

The algorithm first selects the instance to assess. The function findMax(...) in line 1 selects the instance $a^* \in A_C$ with the highest likeliness among those which appear in $T_C$ and for which there is no corresponding instance in $A_U$ (i.e., for which we do not know the user's stance yet). If more than one assertion have the same highest likeliness, one of them is randomly selected. Then, the algorithm moves along the branch connecting $a^*$ to the root in the assessment tree $T_C$ (lines 2-6), until it finds an instance for which the parent already exists in $A_U$ (in the worst case, moving up to the root, which by definition has an instance in $A_U$). To clarify the concept, let us assume that $a^*$ is GB_GREENTEA_EATING_BREAKFAST, but the robot does not yet know whether the user, Mrs. Smith, drinks tea at breakfast, let alone whether she has breakfast at all. The algorithm traces the missing information back towards the root, until it finds the first one to assess (e.g., whether Mrs. Smith has breakfast). This way of proceeding allows for investigating if Mrs. Smith has breakfast (a node closer to the root in the assessment tree) before discussing her breakfast preferences (nodes farther from the root). Moreover, it allows for pruning the branches departing from GB_EATING_BREAKFAST in case it has been assessed

---

[10] $T_C$ is built from $A_C$ using OWL-2 APIs. In principle, building $T_C$ as a separate structure is not required, since the tree–like structure of $A_C$ can be directly explored using OWL-2 APIs. From this perspective, Algorithm 1 shall be interpreted as providing information about how $A_C$ is explored by Algorithms 2 and 3.

**Fig. 14** The assessment tree corresponding to breakfast habits—British culture-specific (GB prefix)
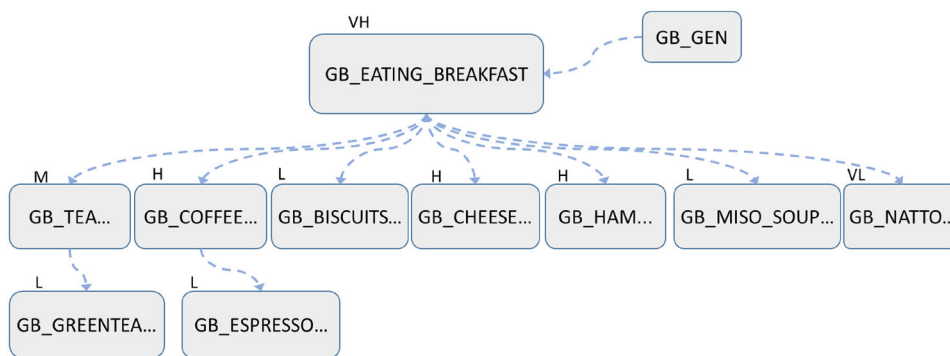


that Mrs Smith does not have breakfast (so that they will not be considered in subsequent iterations of the algorithm). Once the instance $a^*$ to assess has been identified, the routine assess() (line 7) allows for acquiring evidence about it from the user. In the current implementation of the algorithm, the assessment consists in verbally asking the corresponding question to the user (or one among the questions). Then, a new instance in the person-specific ABox is created, to store the newly acquired information (line 8), and $T_C$ is updated (line 9), e.g., eventually pruning portions of it. Sentences stored in the data property hasPositiveSentence are used by the robot to talk about things that it already knows about the user, in order to provide context to its questions, keep the interest of the person and show attentiveness to her values, preferences, beliefs, etc. For instance, if the robot knows that Mrs. Smith usually has breakfast but it does not know if she has tea for breakfast, the sequence of sentences/questions might be the following: hasPositiveSentence="Having a healthy breakfast is important"; hasQuestion="Do you usually have a cup of tea for breakfast?" (since the instance DS_EATING_BREAKFAST exists in the ABox, whereas the instance DS_TEA_EATING_BREAKFAST does not).

## 3.5 Bayesian Adaptation

The assessment and adaptation method discussed in the previous Section has two main limitations:

– it assumes the national culture of the user to be one and a priori fixed;
– it assumes the likeliness of instances to be uncorrelated from each other.

To better clarify the second issue, let us again consider the breakfast habits of different cultures across the world. In many cases, ham and cheese go together: either they are both common options for breakfast, as it happens in Germany, or they are both uncommon, as it happens in Italy. The likeliness of one, thus, can be correlated to the likeliness of the other. Correlations can be found even between instances very

---

**Algorithm 3** Initialization (Bayesian Adaptation)

**Require:** $\mathcal{T} = \{T_C\}$, C $= 1 \dots N$
**Ensure:** $B$
1: init$(B, T_1)$
2: **for all** $a^* \in B$ **do**
3:     $L(a^*) \leftarrow$ getLikeliness$(\mathcal{T}, a^*)$
4:     $CPT(a^*) \leftarrow$ computeCPT$(B, L(a^*))$
5:     update$(CPT(a^*), a^*, B)$
6: **end for**

---

far from each other; for example, one might find that, for a national culture, ham and cheese are common for breakfast and Pentecost Monday is celebrated as a national holiday: if this were true, updating the likeliness of one instance after having acquired evidence about the other might speed up the assessment and adaptation process.

Modelling and managing such correlations requires probabilistic reasoning over the likeliness values, which is a capability that standard ontologies do not have. To this purpose, beside approaches aiming at extending the ontology itself with mechanisms for dealing with probability, such as PR-OWL [7], a large corpus of literature relies on complementing a standard ontology with a Bayesian Network which takes care of the probabilistic reasoning [1,10].

We follow the latter philosophy, and associate to the ontology a Bayesian Network[11] built starting from the assessment trees of all cultures of relevance. As an example, the Bayesian Network in Fig. 15 is built starting from three assessment trees with identical structure as the one shown in Fig. 14, for the Italian, German, and Japanese cultures.

Algorithm 3 shows how the Bayesian Network $B$ is built starting from a set of $N$ assessment trees $\mathcal{T} = \{T_C\}$, with $C = 1 \dots N$. First, the nodes and links of $B$ are built to mirror the structure of one tree in the set $\mathcal{T}$, say $T_1$ (all $T_C$ have the same structure), but the root node of $B$ has a link towards all other nodes (line 1). Then, for each node $a^*$ in $B$, the $N$ likeliness values $L(a^*)$ of the assessment trees $T_C$ (line 3) are used to build the corresponding Conditional Probability

---

**Fig. 15** The Bayesian Network corresponding to breakfast habits—Italian, German, Japanese. Yellow filling denotes culture-specific likeliness values imported from the culture-specific ABox layers
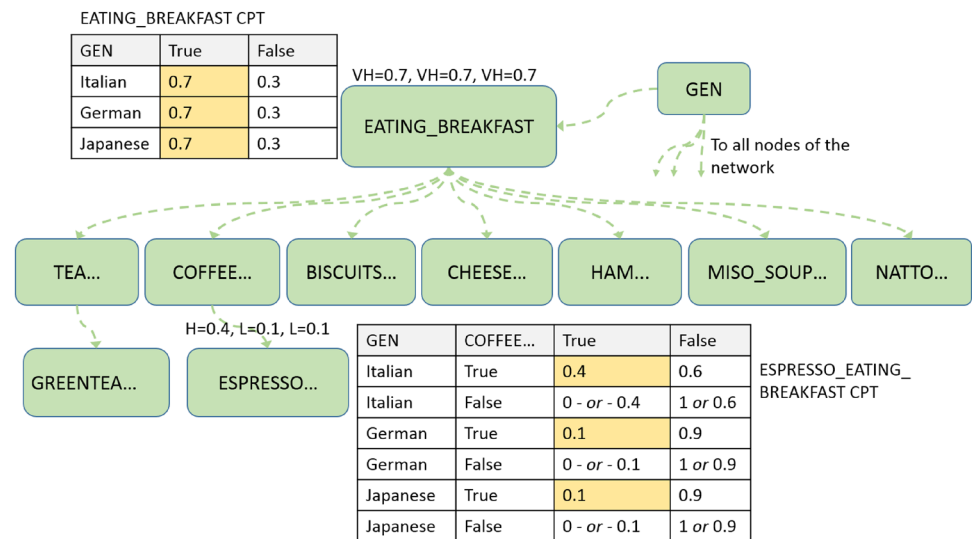


Table (CPT) (line 4) and update the Bayesian Network with the CPT (line 5).

The differences between the Bayesian Network and the assessment trees deserve a detailed analysis.

Firstly, the network does not correspond to the assessment tree of one specific culture (e.g., German or Italian) but it rather stores information about all the cultures taken into account. As an example, the Bayesian Network in Fig. 15 integrates culture-specific knowledge about the Italian, German, and Japanese cultures (notice the missing national prefix).

As a consequence, the root node, unlike all other nodes, is not binary, as it considers all the cultures taken into account (e.g., Italian, German, Japanese). This fact allows for dealing with the first issue listed above: the a priori probability of the root node (i.e., of the user's culture) can be initialized as uniformly distributed over all possible background cultures, or on the basis of available knowledge about the person. For example the a priori probability of the root node for an Italian woman living in Alto-Adige, a German-speaking province in the north of Italy with strong ties with the Austrian culture, might be set as: $P(GEN = Italian) = 0.7$, $P(GEN = German) = 0.3$, $P(GEN = Japanese) = 0$.

Secondly, in the Bayesian Network the node GEN is predecessor of all the nodes of the network. This directly comes from the definition of culture-specific likeliness (see Definition 2), which relies on probabilities directly conditioned by the user's culture. The Bayesian Network makes this dependency explicit.

Lastly, each node of the Bayesian Network is associated with a CPT (two of them are shown in Fig. 15) that represents the probability of the node conditioned by all its parents in the network.

Filling the CPT, the task of function computeCPT(...) in line 4 of Algorithm 3 is not trivial.

Consider the EATING_BREAKFAST CPT of Fig. 15, which contains the values for $P(EATING\_BREAKFAST = T|GEN = Italian)$, $P(EATING\_BREAKFAST = T|GEN = German)$ and $P(EATING\_BREAKFAST = T|GEN = Japanese)$: these values corresponds to the $N = 3$ likeliness values that are stored in the corresponding culture-specific ABox layers (VH, VH, VH in the Figure). The missing entries of the CPT are straightforwardly computed on the basis of the available ones.

However, in the case of nodes (e.g., ESPRESSO_EATING_BREAKFAST) that are conditioned both by the root node GEN and by their immediate predecessor in $T_C$ (e.g., COFFEE_EATING_BREAKFAST), the available likeliness values are not sufficient to define all the entries of the table. The rationale that we propose to address this problem (and that we have adopted in the experiments described in Sect. 4) is as follows. For the CPT entries which assume the immediate predecessor in $T_C$ to be TRUE, the CPT entry is set equal to the culture-specific likeliness, i.e., we assume that the immediate predecessor in $T_C$ has no impact on the probability of the node (see the yellow cells in the CPT of node ESPRESSO_EATING_BREAKFAST). For the CPT entries which assume the immediate predecessor to be FALSE (see rows 2, 4, and 6 in the CPT of node ESPRESSO_EATING_BREAKFAST), we propose two options: (i) *with pruning*, i.e., the CPT entry is set to 0 (which means, for example, that we deem it impossible for a person who does not drink coffee for breakfast to have an espresso for breakfast); (ii) *without pruning*, i.e., the CPT entry is set equal to the culture-specific likeliness, i.e., we again assume that the immediate predecessor has no impact on the probability of the node. As the experiments show, the latter option is better suited for assessment methods relying on direct interaction with the user (thus subject to the logical inconsistencies typ-

ical of natural dialogues), while the former is better suited for assessment methods which rely on sensory data.

Figure 16 shows how the *initialization* and *assessment* phases become when Bayesian adaptation is enabled. The *initialization* phase is devoted to the creation and filling of the Bayesian Network, as outlined in Algorithm 3. The online *Bayesian assessment* phase follows the same steps sketched in Algorithm 2 to assess the personal traits of the user and update the person-specific ABox accordingly. More precisely, once a new evidence is acquired (line 7), it is incorporated in the Bayesian Network (line 8), which results in an update of the posterior probabilities of all other still-unknown nodes, including the root. By copying the new posterior probabilities in the assessment tree, the update ultimately impacts the order in which instances in the ABox are assessed (line 1).

This feature of Bayesian Networks allows for tackling the second issue discussed at the beginning of the Section. A-priori known correlations among instances can be modelled by adding links between nodes and defining CPTs accordingly. In the case of ham, cheese and Pentecost Monday, the correlation is due to the fact that the nodes have a common cause (i.e., they are all valid for a given "national culture of the user" GEN) rather than a causal relationship among them. This case is straightforwardly mapped onto the Bayesian Network, where the probabilities of nodes with a common predecessor are dependent given that their predecessor is not completely known (i.e., $P(HAM\_EATING...|PENTECOST...) \neq P(HAM\_EATING...)$).[12] This situation is very likely to hold in our application, as persons very rarely completely match their background culture (think of the Italian woman living close to Austria, whose culture might be a personal mixture of the Italian and German cultures and for which we might set a non-null probability both for $P(GEN = Italian)$ and $P(GEN = German)$).

## 4 Experimental Evaluation

**Hypothesis** Given a user self-declared as belonging to nationality $G$ and given national-level, culture-specific knowledge about:

- nationality $G$;
- a nationality $A$ geographically close to $G$;
- a nationality $B$ geographically far from $G$;

then, the use of the representation proposed in the previous Sections allows for speeding-up the acquisition of person-

specific knowledge starting from national-level, culture-specific knowledge. Concretely, we postulate that the proposed representation and methods allow for increasing the number of correct "educated guesses" made by the robot, and therefore to identify all instances which hold true for the person with a smaller amount of questions asked.

In particular, we postulate that:

**H1** asking questions using the *likeliness-driven* assessment algorithm described in Sect. 3.4 allows for acquiring person-specific information faster than using a *random* assessment algorithm, which asks questions in a random order;

**H2** asking questions using the likeliness-driven assessment algorithm with *Bayesian adaptation* described in Sect. 3.5 allows for a further speed-up with respect to the *likeliness-driven* assessment algorithm;

**H3** when enabling the *Bayesian adaptation*, the assessment algorithm is able to converge towards the nationality $G$ self-declared by the user, regardless of how it has been initialized.

In the preliminary evaluation reported here we consider two situations:

- the user nationality $G$ is Italian, and we set $A$ as German and $B$ as Japanese;
- the user nationality $G$ is German, and we set $A$ as Italian and $B$ as Japanese.

The Japanese national-level cultural knowledge is considered in the experiments to evaluate hypothesis H3, to allow for assessing the algorithm's performance when initialized with national-level cultural information possibly very far from the user's stances.

### 4.1 Materials and Methods

As the above hypotheses suggest, the proposed framework for the encoding and acquisition of cultural knowledge is novel in itself, independently from its use by a robot in an assistive scenario, and thus requires to be evaluated before and independently from its deployment. Planned future tests will be aimed at assessing the effectiveness, naturalness and cultural competence of an assistive robot equipped with the proposed framework [5]. For a preliminary validation of the proposed approach we have chosen a subset of conversation subject matters (as described in Sect. 3.1) and formulated a list of questions related to these selected topics. In the specific case, five conversation areas have been selected: AttitudeTowardsSport, AttitudeTowardsHolidays, AttitudeTowardsOtherPeople

---

[12] On the opposite, two nodes with a common predecessor are conditionally independent given that their predecessor is known, i.e., $P(HAM\_EATING...|PENTECOST..., GEN) = P(HAM\_EATING...|GEN)$.

**Fig. 16** The offline *initialization* and online *likeliness-driven assessment* phases, which, respectively, create and update the person-specific ABox layer, in the case Bayesian adaptation is enabled
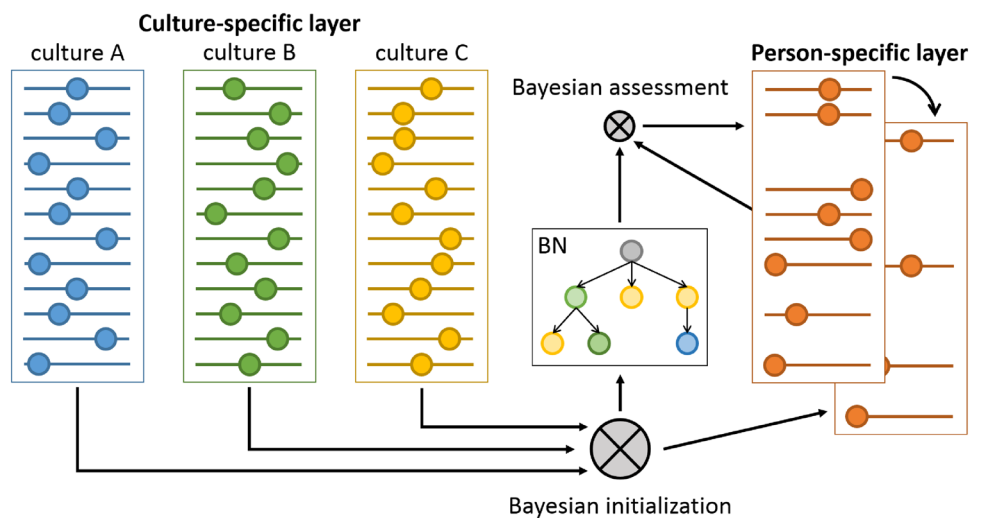


and AttitudeTowardsEating, with a specific focus on the sub-topic EatingBreakfast, for a total of 122 questions. For each question, possible answers are *yes*, *rather yes than no*, *rather no than yes* and *no*.

Questions have been prepared by the authors, after performing a qualitative and quantitative analysis of a high number of documents available online, describing the typical habits and preferences of Italian, German and Japanese persons related to the conversation subject matters mentioned above. Some of the questions capture habits of a specific culture (e.g., attitude towards specific national holidays, or towards typical breakfast food), and for all questions the probability of getting a positive answer from an Italian, German or Japanese user (the *culture-specific likeliness* described in Sect. 3) has been estimated. For example, an Italian user has a low probability of drinking green tea during breakfast (and thus the corresponding instance in the culture-specific ABox has a low likeliness), while he has a high probability of shaking hands while introducing himself (high likeliness).

The list of questions has been made available online and shared mainly by means of social networks;[13] the respondent is requested to provide information about his city of residence, nationality, age and gender, while remaining anonymous. At the end of the experiments, we could collect the answers of 124 Italian and 35 German (some of them living in Italy) respondents.

Collected answers have been analysed offline, with the following methodology. For each respondent, we:

1. initialize all variants of the Assessment & Adaptation algorithms (i.e., *random* assessment, *likeliness-driven* assessment, likeliness-driven assessment with *Bayesian*

*adaptation*), in accordance with the hypothesis under evaluation;

2. for each variant, identify the instance $a^*$ to be asked (i.e., the "educated guess" that the algorithm considers likely to be true for the user, given its current knowledge about him/her);

3. for each variant, retrieve the respondents' answer to $a^*$, update the algorithm and go back to step 2 until all instances for which the respondent gave a positive answer have been found.

In all cases, a question cannot be asked more than once. Since a core feature of a *culturally competent* personal robot is the ability to quickly identify and adapt to the preferences, habits and needs of its user, an algorithm able to assess them in a short time (i.e., able to obtain a greater number of positive answers in the same amount of time) is to be preferred. Thus, considering the answers *yes* and *rather yes than no* as positive answers, the number of questions required by an algorithm to obtain 50%, 60%, 70%, 80%, 90% and 100% of positive answers have been considered as the benchmarking parameters. Moreover, since Algorithm 2 (with and without Bayesian adaptation) chooses randomly between assertions $a$ with the same likeliness, each set of user's answer has been considered for 20 runs and the results averaged.

The *random* assessment algorithm requires no initialization phase and randomly selects the question to pose among all available questions (i.e., all the questions that have not been asked already).

As described in Sect. 3.4, the *likeliness-driven* assessment algorithm requires the setup of the selected culture-specific ABox layer and the initialization procedure described in Algorithm 1 to create the assessment tree. Two variants of the algorithm have been considered: in the variant *without pruning* a negative answer to a parent question in the assess-

---

[13] German version: https://tinyurl.com/ybvr4xeo; Italian version: https://tinyurl.com/y8b3zuub.

ment tree has no impact on the *children* questions; in the variant *with pruning* a negative answer to a parent question prunes the underlying branches of the assessment tree (i.e., if the user answers *no* or *rather no than yes* to the question "Do you usually have breakfast?", all questions related to breakfast habits and preferences will be ignored).

The likeliness-driven assessment with *Bayesian adaptation* requires the setup of the culture-specific ABox layer for all national cultures of relevance and the execution of the initialization procedure described in Sect. 3.5 to create the Bayesian Network and the associated assessment tree. In our experiments, we considered the national-level cultures of Italy, Germany and Japan. The Bayesian Network is initialized by setting the a priori probability of the node GEN as distributed over the three available national cultures (e.g., $P(\text{Italian}) = 0.8$, $P(\text{German}) = 0.1$, $P(\text{Japanese}) = 0.1$), and questions are selected with the rationale of Algorithm 2, as discussed in Sect. 3.5. The answers *yes* and *rather yes than no* are interpreted as *yes* findings and deterministically set the evidence of assertion $a^*$ as $P(a^*) = 1$ (and, symmetrically, *no* and *rather no than yes* are interpreted as *no* findings and set the evidence of $a^*$ as $P(a^*) = 0$). Again, two variants have been considered: in the variant *without pruning* all nodes are connected only to the GEN node and a negative answer to a parent question in the assessment tree has no impact onto the *children* questions (see the Conditional Probability Table shown in Fig. 17a); in the variant *with pruning* all nodes are connected both to the GEN node and to their immediate predecessor in the assessment tree and a negative answer to a parent question in the assessment tree prunes the underlying branches of the assessment tree (see the Conditional Probability Table shown in Fig. 17b).

Beside this deterministic interpretation of the collected evidence, we have also tested a *probabilistic* variant of the likeliness-driven assessment with *Bayesian adaptation*, for which *yes* produces an evidence $P(a^*) = 0.8$, *rather yes than no* produces $P(a^*) = 0.6$, *rather no than yes* produces $P(a^*) = 0.4$, and *no* produces $P(a^*) = 0.2$. In this case pruning never happens, but the user's answers still have a direct influence on the a posteriori probability of the immediate successors of $a^*$ (e.g., if the user answers *rather yes than no*, *rather no than yes* or *no* to the question "Do you usually have breakfast?", then all probabilities related to having coffee during breakfast will be accordingly reduced).

## 4.2 Results and Discussion

Figure 18 compares the performance of the *likeliness-driven without pruning*, *likeliness-driven with pruning*, *Bayesian adaptation without pruning* and *Bayesian adaptation with pruning* assessment algorithms, using the *random* assessment algorithm as a reference, to test hypotheses H1 and H2. The two graphs refer, respectively, to the performance with

COFFEE_EATING_BREAKFAST CPT

| GEN | True | False |
|---|---|---|
| Italian | 0.7 | 0.3 |
| German | 0.4 | 0.6 |
| Japanese | 0.1 | 0.9 |

**(a)** Bayesian Adaptation

COFFEE_EATING_BREAKFAST CPT

| GEN | EATING_BREAKFAST | True | False |
|---|---|---|---|
| Italian | True | 0.7 | 0.3 |
| Italian | False | 0 | 1 |
| German | True | 0.4 | 0.6 |
| German | False | 0 | 1 |
| Japanese | True | 0.1 | 0.9 |
| Japanese | False | 0 | 1 |

**(b)** Bayesian Adaptation with pruning

**Fig. 17** Conditional Probability Tables related to the node COFFEE_EATING_BREAKFAST: **a** in the *Bayesian adaptation without pruning* variant the node is connected only to the GEN node, **b** while in the *Bayesian adaptation with pruning* variant the node is connected also to its immediate predecessor EATING_BREAKFAST

Italian subjects (a) and with German subjects (b), and they are averaged over 20 runs and over the total number of subjects per group. In the graphs, the groups of columns denote the ratio between the number of questions required to reach a certain percentage of positive answers (50%, 60%, 70%, 80%, 90% and 100%, from left to right) by each proposed algorithm, with respect to the *random* assessment algorithm. As an example, the leftmost yellow column states that the *Bayesian adaptation with pruning* assessment algorithm discovers 50% of all positive answers with less than half of the questions required by the *random* assessment algorithm, while the rightmost yellow column states that the same algorithm needs approximately 60% of the number of question required by the *random* assessment algorithm to discover 100% of the positive answers.

As the Figure shows, the *likeliness-driven* assessment algorithm seems to be consistently better than the *random* one (thus supporting hypothesis H1), and the *Bayesian adaptation* assessment algorithm seems to be consistently better than the the *likeliness-driven* one (thus supporting hypothesis H2).

Further considerations arise.

– Independently from the assessment algorithm adopted, the variant *with pruning* performs better than the variant *without pruning*, since all questions that are *children*
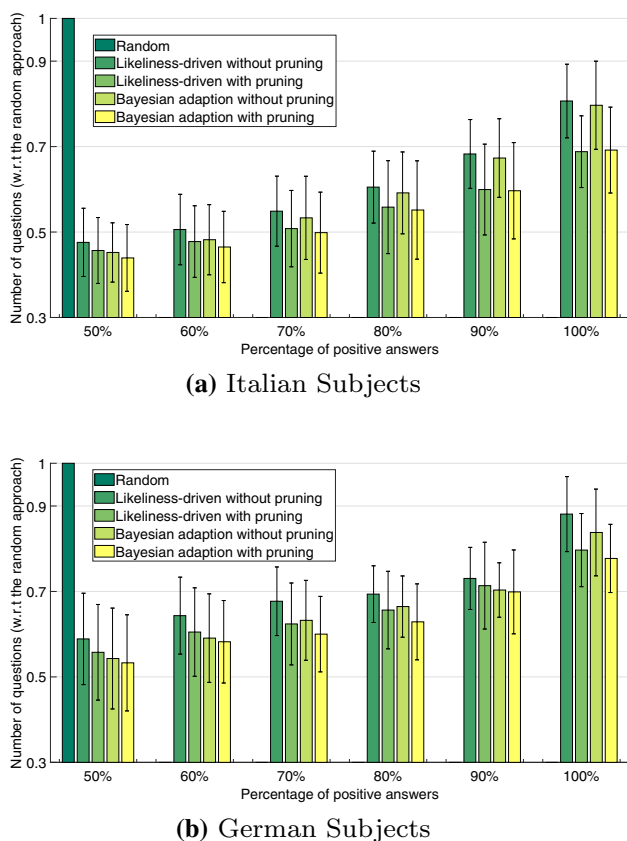
**(a)** Italian Subjects



**(b)** German Subjects

**Fig. 18** Comparison of the performance of the *Likeliness-driven without pruning*, *Likeliness-driven with pruning*, *Bayesian adaptation without pruning* and *Bayesian adaptation with pruning* assessment algorithms, using the *Random* assessment algorithm as a reference. Shorter columns denote better performance

of a question with a negative answer will be not asked. However, since people are often inconsistent in their conversations, these algorithms do not guarantee to obtain the totality of positive answers: in particular, only for 35% of Italian subjects and 26% of German subjects to prune branches does not result in a loss of some positive answers. The two rightmost columns of the variants *with pruning* refer to these subjects only. However, for most of the tests (94% of Italian and German subjects) the variants *with pruning* discover at least 80% of the positive answers.

– The variants of the *Bayesian adaptation* assessment algorithm perform slightly better than the corresponding ones of the *likeliness-driven* assessment algorithm. Indeed, even if the Bayesian Network is initialized with a priori probabilities corresponding to the user nationality, enabling the Bayesian adaptation allows for taking into account the fact that the user may give answers that better match with a different nationality. The advantages brought by this rationale are especially evident with the German subjects, since many of them have lived for a

long time in Italy and this fact has an influence on their answers.

The effects of initializing the GEN node with different probability distributions over the three available national cultures (addressed by hypothesis H3) are shown in Fig. 19. Specifically, the a priori probabilities have been initialized

– with a uniform distribution: P(GEN = Italian) = P(GEN = German)=P(GEN = Japanese) (*no knowledge about user nationality* in the Figure);
– by setting the maximum value corresponding to the nationality *G* declared by the subject in the questionnaire, i.e., for a person that self-declares to be Italian P(GEN = Italian) = 0.8 (*nationality declared by the user* in the Figure);
– by setting the maximum value corresponding to a different nationality A which is geographically close to the one declared by the user, i.e., for a person that self-declares to be Italian P(GEN = German) = 0.8 (*user nationality: case A* in the Figure);
– by setting the maximum value corresponding to a different nationality B which is geographically distant to the one declared by the user, i.e., for a person that self-declares to be Italian, or German P(GEN = Japanese) = 0.8 (*user nationality: case B* in the Figure).

Results show that in all cases there is a clear improvement with respect to the *random* assessment algorithm, even when the Bayesian Network is initialized with a nationality (and therefore, likeliness values) very far from the one declared by the subject (case B). Moreover, the posterior probability of the GEN node (which captures the user's culture), in more than 85% of the tests converges to the nationality declared by the user after sufficient evidence has been collected (thus providing preliminary support to hypothesis H3).

Specifically, for the Italian subjects, initializing the GEN node with the correct nationality of the subjects guarantees better results (see Fig. 19a), while for the German subjects differences in performances between the algorithms are less relevant (see Fig. 19b). To assess whether this is due to the mixed cultural background of the German subjects described before (i.e. many German subjects are living in Italy), Fig. 19c only considers the German subjects who live in Italy. In accordance with our hypothesis, in this case the initialization of the Bayesian Network with the Italian nationality (case A) generally gives better results, as their habits and preferences tend to be more coherent with the culture-specific Italian ABox (e.g., they tend not to have ham and cheese for breakfast).
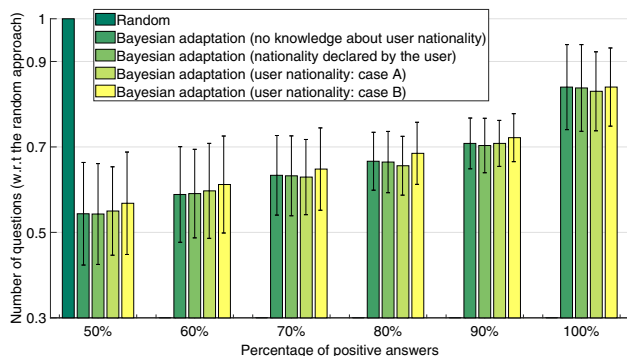
Figure 20 compares the performance of the *Bayesian adaptation without pruning*, *Bayesian adaptation with prun-*
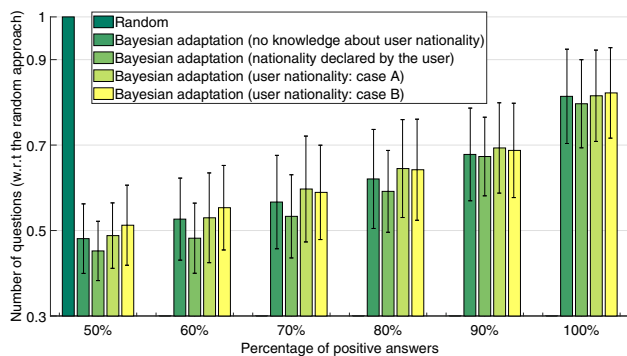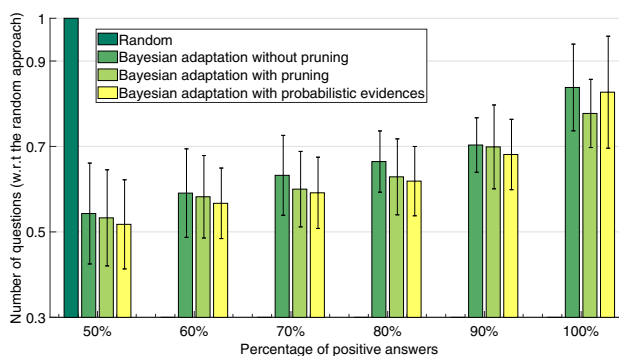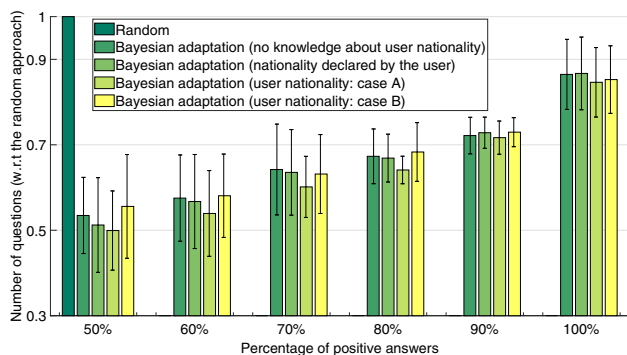
**(a)** Italian Subjects
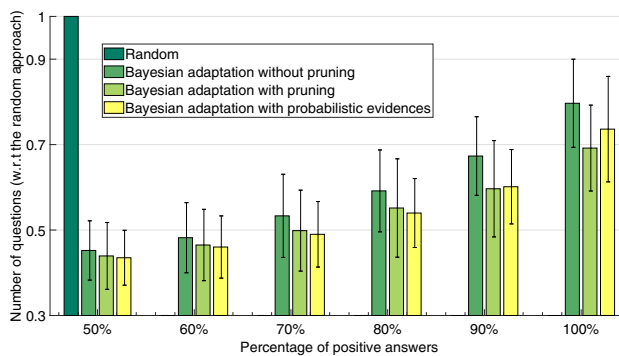


**(b)** German Subjects



**(c)** German Subjects living in Italy

**Fig. 19** Comparison of the performance of the *Bayesian adaptation* assessment algorithm, for different initialization strategies of the a priori probability of the node GEN, using the *Random* assessment algorithm as a reference. Shorter columns denote better performance



**(a)** Italian Subjects



**(b)** German Subjects

**Fig. 20** Comparison of the performance of the *Bayesian adaptation without pruning*, *Bayesian adaptation with pruning* and *Bayesian adaptation with probabilistic evidence* assessment algorithms, using the *Random* assessment algorithm as a reference. Shorter columns denote better performance

Indeed, this variant achieves results that are comparable with those of the variant with pruning, but with no loss of positive answers (i.e., it always finds all positive answers). More specifically, for the Italian subjects, the *Bayesian adaptation with probabilistic evidence* assessment algorithm allows for finding 80% of all positive answers with, on average, 4.99 questions less than the *Bayesian adaptation without pruning* variant, and 1.13 questions less than the variant with pruning.

Finally, the effects of applying the proposed algorithms on the different types of subject matters composing the list of questions (AttitudeTowardsEating, EatingBreakfast, AttitudeTowardsSport, AttitudeTowardsHolidays and AttitudeTowardsOtherPeople) have been investigated. As Fig. 21 shows, while for some subject matters there is a great difference between the performance of the proposed assessment algorithms and the *random* assessment one, for others the difference is less evident. Subject matters of the first type are strictly related to the national culture to which the person belongs, (e.g. AttitudeTowardsHolidays shown in Fig. 21b), while subject matters of the second type typi-

*ing* and *Bayesian adaptation with probabilistic evidence* assessment algorithms. In the first two cases, the evidence is deterministic (i.e., 0 for *no*, *rather no than yes* and 1 for *yes*, *rather yes than no*), while in the latter case it corresponds to 0.8, 0.6, 0.4 or 0.2, moving from *yes* to *no*.

The analysis of the results show that, in general, the *Bayesian adaptation with probabilistic evidence* assessment algorithm performs better than the deterministic variants.

cally include elements that are shared among many cultures (e.g. AttitudeTowardsSport, shown in Fig. 21a).

In the latter case the implementation of a Bayesian Network (initialized with the nationality declared by the user, or even with uniform distribution) gives clear advantages with respect to any variant which does not enable Bayesian adaptation. Conversely, for subject matters strictly related to the national culture, the average performance obtained by the *likeliness-driven* and by the *Bayesian adaptation* assessment algorithms are comparable, although the *Bayesian adaptation* assessment algorithm still performs slightly better than the *likeliness-driven* algorithm. Consider as an example the question "Do you celebrate Pentecost Monday?" - a Christian holiday that is a national holiday in Germany but not in Italy, or the question "Do you celebrate the birthday of the Emperor?" - a Japanese holiday. It is very likely that only German people will give a positive answer to the first question and that only Japanese people will give a positive answer to the second one. However, it cannot be excluded a priori that an Italian person could have acquired some habits that are typical of other cultures: as it happened to some of the German respondents of our questionnaire, the user could have lived for some time in Japan and taken the habit of celebrating the birthday of the Emperor. Thus, while the knowledge of the user nationality is key for quickly detecting the main habits and preferences of the person, an adaptive approach gives the possibility to learn in a shorter time also those traits in which the user differs from its national culture.

## 5 Conclusions

This article tackles the problem of endowing robots with a knowledge representation framework allowing for representing cultural information and using it for better managing and adapting to the user's habits, preferences and needs. Drawing inspiration from the scenarios of culturally competent behaviours for robots for elderly care drafted by experts in Transcultural Nursing, we have identified the main requirements for the robot's knowledge representation system, i.e., (i) the ability to store and manage culture-generic knowledge about the context, the robot itself and the grounding of core values; (ii) the ability to store and manage national-level, culture-specific knowledge, that the robot can rely on whenever person-specific information is not available; (iii) the ability to store and manage person-specific knowledge, describing the way in which the cultural identity, preferences and environment of the assisted person shape the appropriate robot behaviours; and (iv) the ability to efficiently integrate person-specific and culture-specific knowledge, by relying on the latter to discover the former.

To fulfil the above requirements, we have proposed a framework which relies on three core elements: (i) a three-
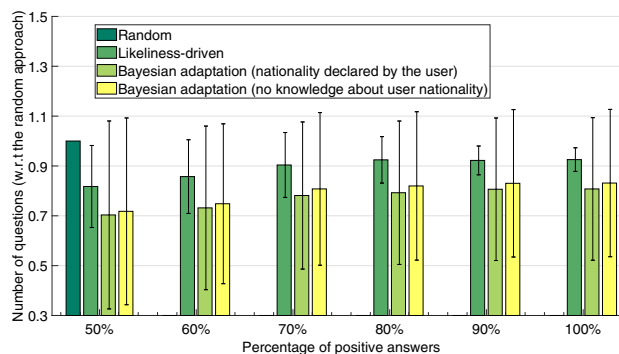


**(a)** AttitudeTowardsSport
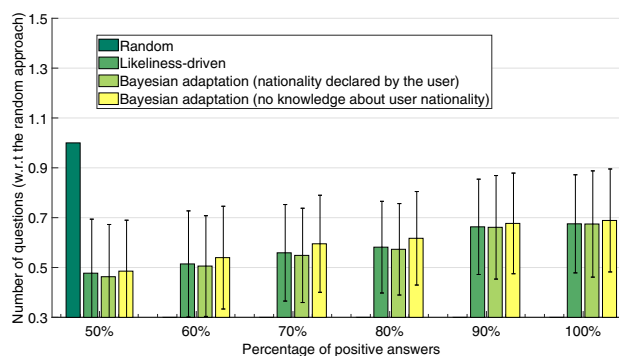


**(b)** AttitudeTowardsHolidays

**Fig. 21** Comparison of the performance of the *likeliness-driven* and *Bayesian adaptation* assessment algorithm, for different initialization strategies of the a priori probability of the node GEN, over different types of conversation subject matters. The *Random* assessment algorithm is taken as a reference. Shorter columns denote better performance. The data refer to the Italian subjects only

layered ontology for storing all concepts of relevance, national-level information and statistics, and person-specific information and preferences; (ii) an algorithm for the acquisition of person-specific knowledge, driven by national-level knowledge (*likeliness-driven* assessment algorithm with its variants), and (iii) a Bayesian Network for speeding up the adaptation by propagating the effects of acquiring one person-specific information onto interconnected concepts (*Bayesian adaptation* assessment algorithm with its variants).

For a preliminary evaluation of the framework we have hypothesised that, given a user that declares herself as belonging to a given cultural group at national level, using the framework with the proposed algorithms can significantly speed-up the acquisition of person-specific knowledge starting from national level knowledge. This hypothesis has been preliminarily validated with 159 Italian and German volunteers by asking questions on 122 habits, attitudes and social norms.

Ongoing work is devoted to relaxing the limitations of acquiring knowledge only through dialog, but rather using the robot's onboard sensors for culturally-competent object and scene recognition. To this end, we are exploring the use of online vision services, which have the advantage of relying on huge training sets continuously updated and maintained [22].

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

1. Agarwal P, Verma R, Mallik A (2016) Ontology based disease diagnosis system with probabilistic inference. In: 1st India international conference on information processing (IICIP), IEEE, pp 1–5
2. Andrist S, Ziadee M, Boukaram H, Mutlu B, Sakr M (2015) Effects of culture on the credibility of robot speech: A comparison between english and arabic. In: HRI, pp 157–164
3. Baader F, Calvanese D, McGuinness D, Nardi D, Patel-Schneider P (2003) The description logic handbook: theory, implementations and applications. Cambridge University Press, New York
4. Bateman CM (2007) Game writing: Narrative skills for videogames. Charles River Media Independence
5. Bruno B, Chong NY, Kamide H, Kanoria S, Lee J, Lim Y, Pandey AK, Papadopoulos C, Papadopoulos I, Pecora F, Saffiotti A, Sgorbissa A (2017a) Paving the way for culturally competent robots: a position paper. In: RO-MAN 2017, pp 430–435
6. Bruno B, Mastrogiovanni F, Pecora F, Sgorbissa A, Saffiotti A (2017b) A framework for culture-aware robots based on fuzzy logic. In: IEEE international conference on fuzzy systems (FUZZ-IEEE), IEEE, pp 1–6
7. Carvalho RN, Laskey KB, Costa PC (2017) Pr-owl-a language for defining probabilistic ontologies. Int J Approx Reason 91:56–79
8. Conti D, Di Nuovo S, Buono S, Di Nuovo A (2017) Robots in education and care of children with developmental disabilities: a study on acceptance by experienced and future professionals. Int J Soc Robot 9(1):51–62
9. Dautenhahn K, Woods S, Kaouri C, Walters ML, Koay KL, Werry I (2005) What is a robot companion-friend, assistant or butler? In: IEEE/RSJ international conference on intelligent robots and systems (IROS 2005) IEEE, pp 1192–1197
10. Ding Z, Peng Y (2004) A probabilistic extension to ontology language owl. In: Proceedings of the 37th annual Hawaii international conference on system sciences, IEEE, p 10
11. Eresha G, Häring M, Endrass B, André E, Obaid M (2013) Investigating the influence of culture on proxemic behaviors for humanoid robots. In: RO-MAN 2013, pp 430–435
12. Evers V, Maldonado H, Brodecki T, Hinds P (2008) Relational vs. group self-construal: untangling the role of national culture in HRI. In: HRI 2008, pp 255–262
13. Flandorfer P (2012) Population ageing and socially assistive robots for elderly persons: the importance of sociodemographic factors for user acceptance. Int J Popul Res 2012:13. https://doi.org/10.1155/2012/829835
14. Guarino N, et al (1998) Formal ontology and information systems. In: Proceedings of FOIS, pp 81–97
15. Hall ET (1976) Beyond culture. Anchor, Norwell
16. Hofstede G, Hofstede GJ, Minkov M (1991) Cultures and organizations: software of the mind, vol 2. McGraw-Hill, New York city
17. IEEE Standards Association (2018) 1872.1-robot task representation. http://standards.ieee.org/develop/project/1872.1.html. Accessed 18 Jan 2019
18. Joosse MP, Poppe RW, Lohse M, Evers V (2014) Cultural differences in how an engagement-seeking robot should approach a group of people. In: CABS 2014, pp 121–130
19. Köckemann U (2016) Constraint-based methods for human-aware planning. Ph.D. thesis, Örebro university
20. Lugrin B, Frommel J, André E (2015) Modeling and evaluating a bayesian network of culture-dependent behaviors. In: Culture computing 2015, pp 33–40
21. Marios Vasiliou R, Christiana Kouta R, Vasilios Raftopoulos R (2013) The use of the cultural competence assessment tool (ccatool) in community nurses: the pilot study and test-retest reliability. Int J Caring Sci 6(1):44
22. Menicatti R, Bruno B, Sgorbissa A (2017) Modelling the influence of cultural information on vision-based human home activity recognition. In: 2017 14th international conference on ubiquitous robots and ambient intelligence (URAI), pp 32–38
23. Niles I, Pease A (2001) Towards a standard upper ontology. In: Proceedings of the international conference on formal ontology in information systems, ACM, pp 2–9
24. Nomura T, Suzuki T, Kanda T, Han J, Shin N, Burke J, Kato K (2008) What people assume about humanoid and animal-type robots: cross-cultural analysis between japan, korea, and the united states. Int J Humanoid Robot 5(01):25–46
25. Papadopoulos I (2006a) The papadopoulos, tilki and taylor model of developing cultural competence. Transcultural health and social care: development of culturally competent practitioners, pp 7–24
26. Papadopoulos I (2006b) Transcultural health and social care: development of culturally competent practitioners. Elsevier, Amsterdam
27. Patompak P, Jeong S, Nilkhamhang I, Chong NY (2017) Learning social relations for culture aware interaction. In: 2017 14th international conference on ubiquitous robots and ambient intelligence (URAI), pp 26–31
28. Rehm M, Bee N, Endrass B, Wissner M, André E (2007) Too close for comfort?: adapting to the user's cultural background. In: HCM 2007, pp 85–94
29. Robinson H, MacDonald B, Broadbent E (2014) The role of healthcare robots for older people at home: a review. Int J Socl Robot 6(4):575–591
30. Ryan JO, Mateas M, Wardrip-Fruin N (2016) A lightweight videogame dialogue manager. In: DiGRA/FDG
31. Torta E, Cuijpers RH, Juola JF, van der Pol D (2011) Design of robust robotic proxemic behaviour. In: ICSR 2011, pp 21–30
32. Trovato G, Zecca M, Do M, Terlemez Ö, Kuramochi M, Waibel A, Asfour T, Takanishi A (2015) A novel greeting selection system for a culture-adaptive humanoid robot. Int J Adv Robot Syst 12:34
33. W3C Owl Working Group and others (2009) OWL 2 web ontology language document overview
34. Wang L, Rau PLP, Evers V, Robinson BK, Hinds P (2010) When in rome: the role of culture & context in adherence to robot recommendations. In: HRI 2010, pp 359–366

**Barbara Bruno** is a post-doc Research Fellow at the University of Genoa, Italy, and Technical Manager of the H2020 project CARESSES. Barbara received the M.Sc. and the Ph.D. in Robotics from the University of Genoa in 2011 and 2015, respectively. She has been Visiting Scholar at Örebro University, Sweden, with the Italian-Swedish project WEARAMI. She has published more than 20 articles in international journals and peer-reviewed international conferences. Her research interests are in Social Robotics, Human-Robot Interaction and Human Activity Recognition.

**Carmine Tommaso Recchiuto** obtained his M.Sc. in Electronic Engineering from the University of Pisa, Italy in 2008 and his Ph.D. in Microsystems from the University of Roma Tor Vergata, Italy in 2015. Since 2014, he is Research Fellow at the Department of Informatics, Bioengineering, Robotics, and System Engineering (DIBRIS), of the University of Genoa, Italy, being involved in the projects PRISMA, focussing his work in the implementation of UAVs in emergency scenarios, and CARESSES (on-going), aimed at the realization of culturally-competent assistive robots. His main interests are in the fields of human-robot interaction, humanoid robotics, motion and path planning.

**Irena Papadopoulos** is the Head of the Research Centre for Transcultural Studies in Health at Middlesex University, UK. Over the last 30 years she has conducted multiple studies focusing on various aspects of transcultural nursing and cultural competence. She is the originator of the study that resulted in the development of the Papadopoulos, Tilki and Taylor Model of Transcultural Nursing and Cultural Competence. She has also developed the Cultural Competence in Action Tool (CCATool) for measuring cultural competence of health professionals working with different patient groups. She has recently coined the term "Culturally Competent Compassion" and developed a new theoretical model and an assessment tool for that. She is the founder of the Intercultural Education for Nurses in Europe (IENE) program which has been funded by the European Commission over the last 10 years. She has authored more than 100 journal articles, numerous research reports, four books on cultural competence and has supervised many doctoral students. She also serves as the elected president of the European Transcultural Nursing Association (ETNA) since 2013.

**Alessandro Saffiotti** is Full Professor of Computer Science at the University of Örebro, Sweden, where he heads the AASS Cognitive Robotic Systems Laboratory. He has been active for more than 25 years in the integration of Artificial Intelligence and Robotics. His research interests encompass artificial intelligence, autonomous robotics, and technology for elderly people. He has published more than 180 papers in international journals and conferences, has organized a large number of international events, and has been involved in more than 15 EU projects and networks. He is on the editorial board of Artificial Intelligence and of the Int Journal of Social Robotics. He is a member of AAAI, a senior member of IEEE, and an EurAI fellow.

**Christina Koulouglioti** holds a Bachelor's Degree in Nursing and a Master's Degree in Health Care Administration from University of Athens, Greece. She completed her PhD at School of Nursing of University of Rochester, USA, and a Post-Doctoral Fellowship at the High Risk Children and Youth Centre of University of Rochester. She has worked as an Emergency Department Nurse in Greece and as a tenure track Assistant Professor at University of Rochester until 2012. The last six years she holds a position as a Senior Research fellow at the Research and Innovation Department of Western Sussex Hospitals NHS Foundation Trust where she manages the local research investigator study portfolio. She has extensive research experience using different quantitative and qualitative methodologies and working with populations across the life span. She is currently seconded to Middlesex University, UK.

**Roberto Menicatti** obtained his BSc in Mechanical Engineering from the University of Genoa, Italy in 2014 and his MSc in Robotics Engineering from the University of Genoa and the École Centrale de Nantes in 2016, as a double degree in the context of EMARO, the European Master in Advanced Robotics. Since 2017, he is Research Assistant at the Department of Informatics, Bioengineering, Robotics, and System Engineering (DIBRIS) of the University of Genoa, being involved in the projects CARESSES (on-going). His main interests are in the fields of natural language understanding and generation, activity recognition and assistive robotics.

**Fulvio Mastrogiovanni** is Assistant Professor at the University of Genoa, Italy, and Affiliate Associate Professor at the Centre for Intelligent Robotics, JAIST, Japan. Fulvio received the MSc and the PhD in Robotics from University of Genoa in 2003 and 2008, respectively. He has served as Management Coordinator for the FP7 project ROBOSKIN, Technical Coordinator for the Italian SHELL project, and PI in other national and EC-funded projects and he is Senior Editor of Intelligent Service Robotics (Springer) and Editor of Robotics and Autonomous Systems (Elsevier). Fulvio has published more than 140 papers in international journals and peer-reviewed international conferences, and one edited book. He is part of the Board of Directors of the Italian Association for Artificial Intelligence. His research interests are in artificial intelligence techniques for Robotics and human-robot interaction.

**Renato Zaccaria** is Full Professor at the University of Genoa, Italy and the Local Coordinator of EMARO+, the European Master in Advanced Robotics. He is the author of more than 140 international scientific articles, of two textbooks, and of two international patents in Robotics. He is the founder of the spin-off company Genova Robot srl, and has been awarded the title of Commendatore by decree of the President of Italy for his technological transfer activity in Robotics-related areas.

**Antonio Sgorbissa** is the Joint Coordinator of the H2020 EU-Japan project CARESSES, Culturally Aware-Robots and Environmental Sensor Systems for Elderly Support. He is Associate Professor at the University of Genoa, Italy, where he teaches Real-Time Operating Systems, Social Robotics and Ambient Intelligence in EMARO+, the European Master in Advanced Robotics. He is PI at University of Genova of the DIONISO project, focussing on Robotics and Wearable Systems for Search & Rescue, and Director of the Inter-university Center for Robotics and Autonomous Systems in Emergency Scenarios (RASES). His research interests are in Knowledge Representation, Context-Awareness, Motion Planning, Wearable and Ubiquitous Robotics. He is the author of more than 150 scientific articles.