# A Formal Definition of Culture

Aliaksandr Birukou, Enrico Blanzieri, Paolo Giorgini, and Fausto Giunchiglia

DISI, University of Trento, via Sommarive 14, Povo, Trento, TN, Italy 38123
{birukou,blanzieri,pgiorgio,fausto}@disi.unitn.it,

**Abstract.** Globalization makes culture no more bound to a geographical area, race or religion as was previously considered in anthropology. With the advent of Web 2.0 it becomes appropriate to speak about the culture of online communities in general, without specific ties to country or nation. Multi-national companies, software developers, scientists need to take into account cultural differences when delivering products to people. The first step in dealing with culture consists in defining and representing culture of the targeted community. AI literature addressed issues of sociality, collaboration, and coordination in agent societies, but did not target the problem of defining and representing culture of a community.

This chapter presents a formal definition of culture of a set of agents. The proposed definition generalizes existing definitions of culture and it is operational in the sense that it can be applied for characterization and comparison of culture(s) existing in various communities. The main focus of this chapter is on the first version of the formalism that does not introduce states. However, representing a snapshot of the culture in a certain moment is the first step towards a more complex formalism that includes states (the work on the latter is ongoing and we present only preliminary definitions here).

**Keywords:** culture, formal definition, agents, communities, Web 2.0, traits, actions, artificial intelligence

## 1 Introduction

The advent of Web 2.0 lead to an explosive growth in the number of applications targeted at communities, e.g. applications supporting social navigation, collaborative editing, bookmarking and tagging. In such applications, culture is no more bound to a geographical area or a religion, as it is usually studied in anthropology. It becomes appropriate to speak about the culture of online communities and such communities in general can not be characterized in terms of race, religion, or country. Rephrasing Axelrod [1], electronic communications allows us to develop patterns of interaction that are chosen rather than imposed by geography. Specific applications such as search engines or e-bookshops and the ways of using them become part of the culture of people. For instance, using Norton Commander file manager or preparing documents in the MS DOS 6.0 operating system, nowadays would be considered unusual to the same extent as lighting one's house with torches. Moreover, in some scenarios we can speak about societies of pure artificial agents, such as web services or programs and their specific culture, e.g. the standards implemented or the set of functionalities used. Human traders and trading agents operating on the same markets together use the same rules and develop common practices which can be referred to as culture that exists in a mixed society composed of different types of agents. All this shows that grasping and representing culture becomes an important problem in computer science. Applications should be developed consistently with the culture of the target community and the notion of culture would provide support for building such applications.

In this chapter we provide a formal definition of culture. Our goal is not to provide a formalism or a reasoning framework *per se*, but, rather, to give an operational definition of culture that can be used for characterizing and describing culture in different scenarios. In particular, we address the problems of development of applications according to the community culture and of characterizing culture of existing communities. We present and formalize a definition of the notion of culture of a set of agents at a moment in time. We define culture as a set of traits that are shared by the set, where traits are characteristics of human societies that are potentially transmitted by non-genetic means and can be owned by an agent (modified from [2]). The requirement of traits being potentially transmitted is needed because transmission is a way of spreading traits, and, consequently, culture, and without transmission it is hard to achieve sharing. The sharing of such traits by the set is required for two reasons: (1) to go from the set of personal traits of an individual to the culture of the set of agents, and (2) to filter out characteristics which only pertain to the set of agents as a whole, but not to individuals. An examples of latter traits is birth rate.

This chapter consolidates and slightly updates the work previously published as a technical report [3] and a workshop paper [4]. We start in Section 2 with the review of the use of concept of culture in the literature. Section 3 presents the first version of the formalism that does not introduce states and does not consider spreading of culture. However, representing a snapshot of the culture of a set of agents in a certain moment is the first step towards a more complex formalism that includes states and models transmission. Section 4 defines some measures on top of the formalism, and an example, illustrating our approach, is presented in Section 5. A preliminary version of the formalism with the states is there presented in Section 6 We discusses related work and limitations of the approach in Section 7, and conclude in Section 8.

## 2 Defining culture

Culture is a slippery and ubiquitous concept. Initially, culture was associated with the notion of civilization tout-court. At the end of the 30s Margaret Mead put in contrast "culture" with "a culture". "Culture means the whole complex of traditional behavior which has been developed by the human race and is successively learned by each generation"([5] cited in [6]). However, specificity of the notion of culture with respect to a given human society was needed in order to study other societies. So the same citation goes on as: “A culture is less precise. It can mean the forms of traditional behavior which are characteristic of a given society, or of a group of societies, or of a certain race, or of certain area, or of a certain period of time” (cited in [6]). As a consequence, in the anthropological literature culture has been introduced as the concept denoting the object of study of cultural anthropology. Other definitions were proposed and they largely vary. However, they seem to converge to the notion that culture is learned [1], it is associated with groups of people and its content includes a wide range of phenomena including norms, values, shared meanings, and patterned ways of behaving [7–12]. In anthropological literature the usefulness of the notion of culture as a scientific tool has been attacked giving rise to the so-called “writing against culture movement” (see Brumann [6] for a reaction against it). The culture as defined in anthropology usually refers to societies defined in national or ethnic terms, however, the concept of culture has been recently used for describing knowledge and behavior of other groups like in the concepts of corporate culture or organizational culture [7, 13, 14]. Moreover, globalization has brought about the problem of interaction of cultures. On the one hand, such interaction leads to blurring boundaries between cultures, while on the other hand it leads to the increasing need of cultural-aware managers and professionals. Recent anthropology textbook definitions take into account the shift in meaning as, for example, in the definition by Peoples and Bailey:

> *Culture is the socially transmitted knowledge and behavior shared by some group of people* (Peoples and Bailey [15, p. 23] cited in [6]).

Earlier authors define culture in the following ways (cited in Brumann [6]):

- Culture ... refers ... to learned, accumulated experience. A culture ... refers to those socially transmitted patterns for behavior characteristic of a particular social group (Keesing [16, p. 68]).
- Culture, or civilization, ... is that complex whole which includes knowledge, belief, art, law, morals, custom, and any other capabilities and habits acquired by man as a member of society (Tylor [17, p. 1]).
- The culture of any society consists of the sum total of ideas, conditioned emotional responses, and patterns of habitual behavior which the members of that society have acquired through instruction or imitation and which they share to a greater or less degree (Linton [18]).
- A culture is the total socially acquired life-way or life-style of a group of people. It consists of the patterned, repetitive ways of thinking, feeling, and acting that are characteristic of the members of a particular society or segment of a society (Harris [19]).

As we can see, definitions agree on the fact that culture consists of something that is shared and/or learned by a group of people, but the content of the culture varies in different definitions. Similarly to Axelrod [1], we see the content of the culture as a set of traits[1], which can refer to behavior, knowledge facts, ideas, beliefs, norms, etc.

## 3 A formal definition of culture. The version without states

Consistently with the AI literature, we define an *agent* as a "physical or virtual entity that can act, perceive its environment (in a partial way) and communicate with others, is autonomous and has skills to achieve its goals and tendencies" [20]. An agent can represent an individual or a collective entity such as an organization, and can have different *cultural traits*, which are characteristics of human societies that are potentially transmitted by non-genetic means and can be owned by an agent. The requirement "can be owned by", which we add to the definition by Mulder [2], means that it is possible for an agent to have a cultural trait. Behavior, beliefs, knowledge are particular kinds of cultural traits.

Let us consider the set of agents $Ag$ and the set of traits $\mathcal{T}$. Given an agent $a \in Ag$ we denote its set of cultural traits with $T_a = \{\tau_i\} \subseteq \mathcal{T}$ and we use the predicate $has(a, \tau)$ to represent the fact that the agent $a$ has a trait $\tau \in T_a$. In the following, we call the set of traits of an individual *the culture of an individual.*

| **trait type** | **traits** |
|---|---|
| knowledge | $Dante\_Alighieri\_wrote\_Divine\_Comedy(DA)$, $cappuccino\_is\_coffee(CI)$, |
| | $latte\_macchiato\_is\_coffee(LM)$, $Meiji\_era\_was\_in\_1868\_1912(ME)$ |
| behavior | $eating\_with\_sticks(ES)$, $eating\_with\_fork(EF)$ |
| norms, rules | $never\_put\_mayonnaise\_on\_pizza(NP)$, |
| | $never\_open\_umbrella\_inside\_building(NO)$ |
| beliefs | $Christianity(Chr)$, $Buddhism(Bud)$ |

**Table 1.** The set of traits $\mathcal{T}$ in Example 1.

**Example 1.** Let $Ag$ in our example be a set of people: Charlie, Pedro, Maria, and Andrea are European citizens, and Toru is from Japan. Let $\mathcal{T}$ be a set of traits of different types, as shown in Table 1. For each trait, we also put its abbreviation (used in the figures in this section) in parentheses.

Table 2 lists the set of traits $\mathcal{T}$, and the sets of traits of the specific agents of $Ag = \{Charlie, Pedro, Toru, Maria, Andrea\}$.

We can write $has(Maria, Dante\_Alighieri\_wrote\_Divine\_Comedy)$, or $has(Charlie, cappuccino\_is\_coffee)$, but not $has(Andrea, eating\_with\_sticks)$. We will use this example as a running example through the section ⊘

Note that we introduce types of traits and use them in the example only for convenience. One might propose a different classification of traits, e.g. putting

[1] Traits are further grouped in *features* in Axelrod's formulation, i.e. each feature can take value from a set of specific traits.

| set | traits |
|---|---|
| $T_{Charlie}$ | $Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $cappuccino\_is\_coffee$, $eating\_with\_sticks$, $eating\_with\_fork$, $never\_put\_mayonnaise\_on\_pizza$, $Buddhism$ |
| $T_{Pedro}$ | $Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $cappuccino\_is\_coffee$, $eating\_with\_fork$, $Christianity$ |
| $T_{Toru}$ | $Meiji_era_was_in_$1868$_$1912, $cappuccino\_is\_coffee$, $eating\_with\_sticks$, $Buddhism$ |
| $T_{Maria}$ | $Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $cappuccino\_is\_coffee$, $eating\_with\_sticks$, $eating\_with\_fork$, $Christianity$ |
| $T_{Andrea}$ | $Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $cappuccino\_is\_coffee$, $eating\_with\_fork$, $Christianity$ |

**Table 2.** Traits of agents in Example 1.

$eating\_with\_fork$ as a norm. We believe that there is no single classification and this suggests that our approach of dealing with generic traits rather than with specific types of cultural content provides certain advantages.

**Definition 1 (sharing).** *For each pair of agents $a_i, a_j \in Ag$ and for each trait $\tau \in \mathcal{T}$, $a_i$ and $a_j$ share the trait $\tau$ iff they both have such a trait:*

$$has(a_i, \tau) \wedge has(a_j, \tau) \leftrightarrow sharing(a_i, a_j, \tau).$$

*Property 1.* Sharing is transitive:

$$sharing(a_i, a_j, \tau) \wedge sharing(a_j, a_k, \tau) \rightarrow sharing(a_i, a_k, \tau).$$

**Example 1 (continued).** In the example, we can write $sharing(Toru, Maria, eating\_with\_sticks)$, or $sharing(Pedro, Andrea, cappuccino\_is\_coffee)$, etc. To avoid giving the complete list of tuples for which *sharing* holds, we represent them as a graph where nodes are agents and labels on each edge denote traits that are shared by the pair of agents connected by the edge, see Figure 1.

We can show a restriction of *sharing* to specific traits, as in Figure 2, which shows how one trait, $cappuccino\_is\_coffee$, is shared by the set of agents. ⊘

Given a set of agents $G \subseteq Ag$ and a set of traits $T_G \subseteq \mathcal{T}$ we define the notions of *weak sharing* and *strong sharing*.

**Definition 2 (weak sharing).** *A set of traits $T_G$ is weakly shared by a set of agents $G$ iff for each trait $\tau \in T_G$ there exists a pair of agents $a_i, a_j \in G$, $a_i \neq a_j$ that share $\tau$.*

**Definition 3 (strong sharing).** *A set of traits $T_G$ is strongly shared by a set of agents $G$ iff each trait $\tau \in T_G$ is shared by all pairs of agents $a_i, a_j \in G$, $a_i \neq a_j$.*

**Example 1 (continued).** Let us consider the set of agents $G = \{Charlie, Toru, Maria, Andrea, Pedro\}$ and two sets of traits: $T_G = \{cappuccino\_is\_coffee,$

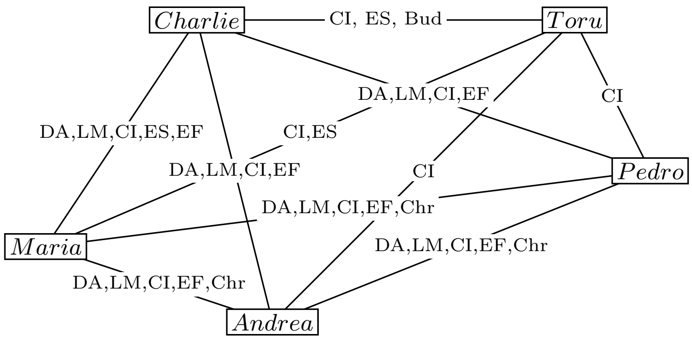


**Fig. 1.** The graph showing for which agents and traits the predicate *sharing* holds in Example 1. The nodes are agents and labels on each edge denote traits that are shared by the pair of agents connected by the edge. For instance, the edge between $Toru$ and $Andrea$ labeled $CI$ means that $sharing(Andrea, Toru, cappuccino\_is\_coffee)$. The traits are abbreviated as in Table 1: $Dante\_Alighieri\_wrote\_Divine\_Comedy$ is abbreviated as $DA$, $latte\_macchiato\_is\_coffee$ as $LM$, $cappuccino\_is\_coffee$ as $CI$, $eating\_with\_sticks$ as $ES$, $eating\_with\_fork$ as $EF$, $Christianity$ as $Chr$, $Buddhism$ as $Bud$.

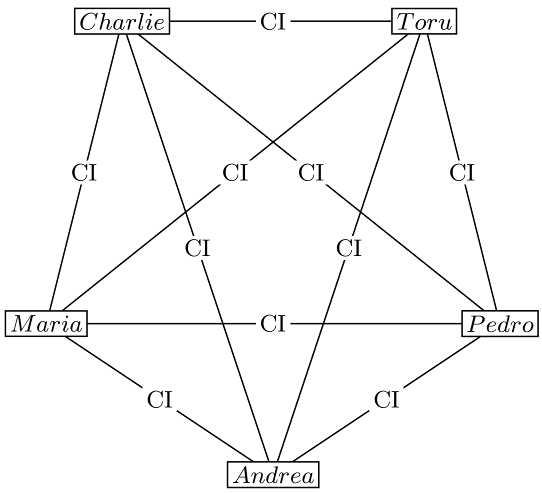


**Fig. 2.** The graph that shows for which agents the *sharing* predicate holds for the $cappuccino\_is\_coffee(CI)$ trait in Example 1.

$eating\_with\_sticks$,$Dante\_Alighieri\_wrote\_Divine\_Comedy\}$, $T'_G = \{cappuccino\_is\_coffee\}$. Using the *sharing* predicate represented in Figure 1, we can see that $T'_G$ is strongly shared by $G$ becaise the $cappuccino\_is\_coffee$ trait is shared by each pair of agents. $T_G$ contains three traits that are shared by at least one pair of agents: e.g., $cappuccino\_is\_coffee$, $eating\_with\_sticks$ shared by $Toru$ and $Charlie$, $Dante\_Alighieri\_wrote\_Divine\_Comedy$ shared by $Charlie$ and $Andrea$. So, $T_G$ is weakly shared by $G$. ⊘

*Property 2.* Strong sharing implies weak sharing.

**Proof.** Strong sharing of a set of traits $T_G$ by a set of agents $G$ means that for each $\tau \in T_G$ all pairs of agents $a_i, a_j \in G$ share $\tau$. Thus, the condition for weak sharing, i.e. existence of one pair of agents $a_i, a_j \in G,\ a_i \neq a_j$ that share $\tau$ is fulfilled. □

Given a set of agents $G \subseteq Ag$ such that $|G| \geq 2$, we introduce the notion of culture of $G$.

**Definition 4 (weak culture of a set of agents).** *A non-empty set of traits* $T_G \subseteq \mathcal{T}$ *is a weak culture of* $G$ *iff*

- *the set* $T_G$ *is weakly shared by* $G$*,*
- *for each agent* $a \in G$ *there exists a trait* $\tau \in T_G$ *such that* $has(a, \tau)$*.*

In other words, Definition 4 says that for a set of agents, a culture is defined as a set of traits weakly shared by agents, and each agent has at least one trait in the culture.

**Definition 5 (strong culture of a set of agents).** *If* $T_G$ *in Definition 4 is also strongly shared then it is* a strong culture of a set of agents.

In the following if we refer to "a culture of a set of agents", we mean "a weak culture of a set of agents".

**Example 1 (continued).** We first give examples of sets of traits that are not culture because one of the conditions for being culture is not satisfied and then give an example of weak and strong cultures.

Let us consider the set $G = \{Pedro, Maria\}$, and the set $T_G = \{eating\_with\_sticks\}$. $T_G$ is not a culture of $G$, because $T_G$ is not weakly shared by $G$.

Let us add $Toru$ in the set. Now, the set $T_{G'} = \{eating\_with\_sticks\}$ is not a culture of $G' = \{Pedro, Maria, Toru\}$ because $Pedro$ does not have the $eating\_with\_sticks$ trait.

On the other hand, the set $T'_{G'} = \{Dante\_Alighieri\_wrote\_Divine\_Comedy$, $eating\_with\_sticks\}$ is a weak culture of $G'$ since the traits $eating\_with\_sticks$ and $Dante\_Alighieri\_wrote\_Divine\_Comedy$ are weakly shared by $G'$ and each agent has at least one trait in $T'_{G'}$ ($Toru$ and $Maria$ have $eating\_with\_sticks$, $Pedro$ has $Dante\_Alighieri\_wrote\_Divine\_Comedy$).

Taking $G'' = \{Toru, Maria\}$, $T_{G''} = \{eating\_with\_sticks\}$ is a strong culture, because it is strongly shared by the set, and each agent has the $eating\_with\_sticks$ trait. ⊘

*Property 3.* Given a set of agents $G \subseteq Ag$ and $T_G$, a culture of $G$, it is possible to find a non-empty set $G_0 \subseteq G$ and a non-empty set $T_{G_0}$ such that $T_{G_0}$ is a strong culture of $G_0$.

**Proof.** If $|G| = 2$ then all traits that are weakly shared are also strongly shared and $T_{G_0} = T_G$ is a strong culture of $G_0 = G$. Otherwise, let us consider $G'_0 = \{a_1, a_2\}$, where $a_1$ and $a_2$ are two agents of $G$ such that $T_{a_1} \cap T_{a_2} \cap T_G \neq \emptyset$ (agents share at least one trait from the culture). The existence of such a pair of agents is guaranteed, because $T_G$ is weakly shared, so for every $\tau \in T_G$ there are at least two agents that share it. Since both $a_1$ and $a_2$ have each trait from $T_{G_0} = T_{a_1} \cap T_{a_2} \cap T_G$, it is strongly shared and the second condition required for a set to be a strong culture is fulfilled. $\square$

## 4 Measures for the comparison of cultures

In this section, we present some measures for characterizing a culture of a set of agents in different socio-cultural settings and for comparing cultures of different sets. This list is not exhaustive, rather, it contains some initial measures, and further extension of this list is a subject of future research.

### 4.1 Measuring culture as a snapshot

**Culture** Let us start from simple measures such as *presence of a specific trait in a culture*. We use an indicator function $I_{has}(\tau, T)$ to say that the trait $\tau$ is present in the culture $T$:

$$I_{has}(\tau, T) = \begin{cases} 1, & \text{if } \tau \in T \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Another example of a simple measure of a culture could be *the number of traits in the culture*, defined as $|T|$, i.e. the dimension of the set of traits $T$.

**Culture of a group** A culture of a group is a product of the individuals belonging to the group. However, different groups can share cultures to some extent. To measure such degree of sharing we adapt the notion of cultural homogeneity introduced by Carley in [21]. Culture in that paper is defined as the distribution of information (ideas, beliefs, concepts, technical knowledge, etc.) across population. In our settings, given a set of agents $G$ and a culture $T_G$ of $G$, the cultural homogeneity is measured by the percentage of possible dyadically shared traits that actually are shared. A trait $\tau$ is shared by a dyad if $sharing(a_i, a_j, \tau)$. The number of possible dyadically shared traits is $\binom{N}{2} \times K$, where $N$ is the number of agents in the set, $|G|$; $K = |T_G|$ is the number of traits in the culture $T_G$.

Thus, *cultural homogeneity* is measured as

$$CH(G,T_G) = \frac{\sum_{i=1}^{N}\sum_{j=i+1}^{N}\sum_{k=1}^{K} I_{sharing}(a_i,a_j,\tau_k)}{\binom{N}{2}\times K}\times 100\%. \quad (2)$$

In this formula, $G = \{a_i\},\ 1 \le i \le N$, $T_G = \{\tau_k\},\ 1 \le k \le K$, and the indicator function $I_{sharing}$ is defined as follows:

$$I_{sharing}(a_i,a_j,\tau_k) = \begin{cases} 1 \text{ , if } sharing(a_i,a_j,\tau_k) \\ 0 \text{ , otherwise.}\end{cases}$$

It is easy to note that the cultural homogeneity takes into account only traits present in the culture, and it does not matter what traits agents of $G$ have besides those contained in the culture $T_G$. To take the traits that are not a part of culture into account, we introduce the notion of group homogeneity. To do this, we need to consider the set of all traits of the group $\bar{T}_G = \cup_{i=1}^{N} a_i,\ \bar{K} = |\bar{T}_G|$. Thus, *group homogeneity* of the group $G$ is measured as

$$GH(G) = \frac{\sum_{i=1}^{N}\sum_{j=i+1}^{N}\sum_{k=1}^{\bar{K}} I_{sharing}(a_i,a_j,\tau_k)}{\binom{N}{2}\times \bar{K}}\times 100\%, \quad (3)$$

where $\tau_k,\ 1 \le k \le \bar{K}$ are from the set $\bar{T}_G$ and the other terms are defined in Equation 2.

**A culture of an individual and a culture of a group** To compare a culture of an individual $a$ and a culture of a group $G$ we introduce the following measures:

- *Common culture (culture overlap)* is the set of traits that is present in both cultures: $CC(T_a,T_G) = T_a \cap T_G$.
- *Culture similarity* is the degree to which two cultures are similar, i.e. how much they have in common: $CS(T_a,T_G) = \frac{|T_a \cap T_G|}{|T_a \cup T_G|} \times 100\%$.
- *Culture fit* is the degree to which one culture fits the other culture: $CF(T_a,T_G) = \frac{|T_a \cap T_G|}{|T_G|} \times 100\%$.
  Note that this measure is not symmetric.

Note that it is possible to extend the notion of culture similarity further if we assume there is a domain-specific function for calculating similarity between traits, i.e. for each pair of traits $\tau_1,\tau_2$ we know the value of $sim(\tau_1,\tau_2)$. Culture similarity can then be defined as $CC(T_a,T_G) = \frac{\sum_{i=1}^{|T_a|}\sum_{j=1}^{|T_G|} sim(\tau_i^a,\tau_j^G)}{|T_a|\times|T_G|} \times 100\%$
This will allow for considering the degree of similarity between different traits, e.g., specifying that trait $eating\_with\_sticks$ is more similar to $eating\_with\_fork$ than to $telling\_DA$.

**A culture of a group and a culture of another group** In order to compare cultures of two sets of agents we can straightforwardly replace the culture of an individual with a culture of another group in the formulas above, thus introducing the following measures:

- Common culture (culture overlap) is the set of traits that is present in both cultures: $CC(T_{G_1}, T_{G_2}) = T_{G_1} \cap T_{G_2}$.
- Culture similarity is the degree to which two cultures are similar, i.e. how much they have in common: $CS(T_{G_1}, T_{G_2}) = \frac{|T_{G_1} \cap T_{G_2}|}{|T_{G_1} \cup T_{G_2}|} \times 100\%$.
- Culture fit is the degree to which one culture fits the other culture: $CF(T_{G_1}, T_{G_2}) = \frac{|T_{G_1} \cap T_{G_2}|}{|T_{G_2}|} \times 100\%$.
  Note that this measure is not symmetric.

| measure | meaning |
|---|---|
| $I_{has}(\tau, T)$ | shows if the trait $\tau$ is present in the culture $T$ |
| $\|T\|$ | the number of elements in the culture $T$ |
| $CH(G, T_G)$ | cultural homogeneity of $G$, i.e. how widely the culture $T_G$ is shared within the group $G$ |
| $GH(G)$ | group homogeneity, i.e. how similar are the sets of traits of agents of $G$ |
| $CC(T_a, T_G)$ $CC(T_{G_1}, T_{G_2})$ | common culture, i.e. the set of traits contained in the culture of an agent $a$ (a group $G_1$) and in the culture $T_G$ ($T_{G_1}$) |
| $CS(T_a, T_G)$ $CS(T_{G_1}, T_{G_2})$ | culture similarity, i.e. how much two cultures have in common |
| $CF(T_a, T_G)$ $CF(T_{G_1}, T_{G_2})$ | culture fit, i.e. the degree to which the culture of $a$ ($G_1$) fits the culture $T_G$ ($T_{G_2}$) |

**Table 3.** Measures of culture as a snapshot.

Table 3 summarizes the presented measures.

### 4.2 Example

Let us see how the described measures apply to Example 1 from Section 3, summarized in Table 2. Considering a set of agents $G = \{Charlie, Toru, Andrea, Maria, Pedro\}$, and a culture $T_G = \{Dante\_Alighieri\_wrote\_Divine\_Comedy(DA), eating\_with\_sticks(ES)\}$:

- $I_{has}(eating\_with\_sticks, T_G) = 1$,
- $I_{has}(eating\_with\_fork, T_G) = 0$,
- $|T_G| = 2$.

To calculate the cultural homogeneity of $G$ we need to calculate the number of traits in the culture $T_G$: $K = 2$ and the number of agents in the set $G$: $N = 5$. With these parameters, $CH(G, T_G)$ is calculated as follows:

$$CH(G, T_G) = \frac{\sum_{i=1}^{5} \sum_{j=i+1}^{5} \sum_{k=1}^{2} I_{sharing}(a_i, a_j, \tau_k)}{\binom{5}{2} \times 2} \times 100\% =$$

$$= \frac{\sum_{i=1}^{5} \sum_{j=i+1}^{5} (I_{sharing}(a_i, a_j, DA) + I_{sharing}(a_i, a_j, ES))}{10 \times 2} \times 100\% =$$

$$\frac{6+3}{20} \times 100\% = 45\%.$$

Proceeding with calculations we get:

- $GH(G) = \frac{41}{120} \times 100\% = 34.17\%$,
- $CC(Pedro, T_G) = \{Dante\_Alighieri\_wrote\_Divine\_Comedy\}$,
- $CS(Pedro, T_G) = 0.125$,
- $CF(Pedro, T_G) = 0.5$,
- $CF(T_G, Pedro) = 0.143$.

## 5 A case study

In this section, we provide a case study that shows how the material presented in this chapter can be applied in the Web 2.0 domain. We first describe the scenario and then show how it can be addressed with our approach.

### 5.1 Scenario description

Let us consider activities related to bibliography management in CiteULike.org, a free online service to organize someone's collection of academic papers. Users of CiteULike are mainly scientists and there are groups dedicated to specific interests. The site allows people to add papers in their personal collections or to the collections of the groups users belong to and to tag those papers. It is also possible to search for the papers using keywords or browse the papers with a specific tag.

Let us suppose that Michael, a user of CiteULike, has some papers about recommendation systems in his bibliography and has tagged them as shown in Table 4[2]. He discovers that there are groups on CiteULike and that there are at least three groups that seem relevant to his research interests: GroupA, GroupB, and GroupC. In the group bibliography, each group has a list of papers tagged

[2] Of course, we present a simplified example here, real users and groups on CiteULike have much more papers in their bibliographies.

| **Michael** | | |
|---|---|---|
| **paperID** | **paper** | **tags** |
| PolyLens | PolyLens: a recommender system for groups of users | recommendation, collaborative filtering |
| TrustInRS | Trust in recommender systems | trust, recommendation |
| GroupLens | GroupLens: An Open Architecture for Collaborative Filtering of Netnews | collaborative filtering, grouplens |
| RefWeb | Referral Web: Combining Social Networks and Collaborative Filtering | collaborative filtering, trust |
| TrustCF | Trust-Aware Collaborative Filtering for Recommender Systems | trust, recommendation |
| **Group A** | | |
| EComRec | E-Commerce Recommendation Applications | collaborative filtering, ecommerce, recommender |
| TechLens | Enhancing digital libraries with TechLens+ | recommender, academic reference |
| GetToKnow | Getting to know you: learning new user preferences in recommender systems | collaborative filtering, recommender |
| GroupLens | Group Lens: An open architecture for collaborative filtering of netnews | collaborative filtering, recommender |
| PolyLens | PolyLens: a recommender system for groups of users | recommendation, collaborative filtering |
| **Group B** | | |
| TechLens | Enhancing Digital Libraries with TechLens+ | collaborative filtering, content based filtering, papers, recommender systems |
| Citations | On the Recommending of Citations for Research Papers | citations, collaborative filtering, personalization, recommender systems |
| Scouts | Scouts, promoters, and connectors: The roles of ratings in nearest-neighbor collaborative filtering | recommender systems, recommendation, collaborative filtering |
| EComRec | E-Commerce Recommendation Applications | collaborative filtering, ecommerce, recommender |
| ContRec | A content-collaborative recommender that exploits WordNet-based user profiles for neighborhood formation | collaborative filtering, concept extraction, concept map, recommender |
| **Group C** | | |
| GroupLens | Group Lens: An open architecture for collaborative filtering of netnews | collaborative filtering, recommender, recommendation |
| VirtCom | Recommending and evaluating choices in a virtual community of use | collaborative filtering, recommender |
| TagCF | Tag-aware recommender systems by fusion of collaborative filtering algorithms | tagging, recommender, collaborative filtering |
| TrustInRS | Trust in recommender systems | trust, recommender, collaborative filtering |
| RefWeb | Referral Web: Combining Social Networks and Collaborative Filtering | collaborative filtering, social network |

**Table 4.** Users and groups in CiteULike.org.

as shown in Table 4. Michael would like to join some group, but he does not have much time to read group feeds, so he would like to choose only one group. How does he decide which group fits more with his interests? The bibliography of a group contains several hundred of items, looking through them will take some time.

Let us assume that all tags are from the same taxonomy and there are no syntactical (e.g., tags *recommendation system*, *recommender systems*, *RS* are replaced with a single tag) and semantical (e.g., tags like *recommendation system*, *adaptive system* correspond to very same concepts in all bibliographies) inconsistencies in the names of papers and tags. Thus, we can represent a group or a user as a set of tags and a set of papers in their bibliography and calculate the degree of the fit between a user and a group as similarity between their sets of tags and papers. Moreover, we can see which papers are common for all three groups, creating for Michael a list of papers to read.

### 5.2 Applying our approach

In our formalism, the users and groups are agents that are represented as a set of traits, which are papers and tags. For each agent, its culture is the set of traits as follows:
$Michael.papers$=$\{PolyLens, TrustInRS, GroupLens, RefWeb, TrustCF\}$
$Michael.tags$=$\{recommendation, collaborative\ filtering, trust, grouplens\}$
$GroupA.papers$=$\{EComRec, TechLens, GetToKnow, GroupLens, PolyLens\}$
$GroupA.tags$=$\{collaborative\ filtering, recommendation, academic\ reference, recommender, ecommerce\}$
$GroupB.papers$=$\{TechLens, Citations, Scouts, EComRec, ContRec\}$
$GroupB.tags$=$\{collaborative\ filtering, content\ based\ filtering, papers, citations, recommender\ systems, personalization, recommendation, ecommerce, recommender, concept\ extraction, concept\ map\}$
$GroupC.papers$=$\{GroupLens, VirtCom, TagCF, TrustInRS, RefWeb\}$
$GroupC.tags$=$\{collaborative\ filtering, recommender, recommendation, tagging, trust, social\ network\}$

Let us select one of the metrics from Section 4, say culture similarity, for determining how close are two cultures. Since the number of distinct papers in $Michael$'s and $GroupA$ bibliographies is eight, the number of common papers is two, the number of distinct tags is seven and the number of common tags is two, the similarity between $Michael$ and $GroupA$, $CS(Michael, GroupA)$ is equal to $0.5 \cdot \frac{2}{8} + 0.5 \cdot \frac{2}{7} = 0.268$. The similarity between $Michael$ and $GroupB$ is $0.5 \cdot \frac{0}{10} + 0.5 \cdot \frac{2}{13} = 0.077$, while the similarity between $Michael$ and $GroupC$ is $0.5 \cdot \frac{3}{7} + 0.5 \cdot \frac{3}{7} = 0.429$. From this simple exercise we can conclude that $Michael$'s research interests, as represented by his bibliography, are closer to $GroupC$. The program realizing such algorithm in real CiteULike.org settings, i.e. with hundreds of groups with thousands of papers, would solve the abovementioned problem of choosing which community to join.

Let us further illustrate how our formalism can be applied to these data. Let us consider each group as an agent and see which traits are shared by the set of agents $\{GroupA, GroupB, GroupC\}$. Papers $EComRec$, $TechLens$, $GroupLens$ and tags *recommender*, *ecommerce*, *recommendation*, *collaborative filtering* are weakly shared by the set and therefore are a culture of the set. Moreover, while there are no strongly shared papers, tags *collaborative filtering*, *recommender*, *recommendation* are strongly shared and therefore are a strong culture of the set.

### 5.3 Discussion

In the case study we calculated the degree of culture similarity between Michael and different groups, and computed a culture of a set of CiteULike groups. Further extending this example, we might take into account not only artifacts such as papers or tags, but also behaviors of users, such as tagging some paper with a specific tag. For instance, using information about authors of the papers and citations, it is possible to consider behaviors such as self-citation and to see if there are communities whose members follow this practice more than an average author. Using information about the publication date and the date of posting the publication in someone's library it is possible to consider behaviors such as "tagging paper before its publication" and see which communities have the practice of dissemination of drafts of the papers.

## 6 A formal definition of culture. The version with states

This section presents a preliminary version of the definition of culture with states. Traits and agents are defined as in Section 3.

To model changes in the set of traits of an agent and consequently, changes in culture, we use the notion of state. We assume that the world can be in different states and the set of traits of the same agent can be different in different states.

Let us consider the set of agents $Ag$, the set of traits $\mathcal{T}$, and the set of states $S$. Given an agent $a \in Ag$ and a state $s \in S$, we denote the set of cultural traits of the agent $a$ in the state $s$ with $T_a(s) = \{\tau_i\} \subseteq \mathcal{T}$ and we use the predicate $has(a, \tau_i, s)$ to represent the fact that the agent $a$ has a trait $\tau_i \in T_a(s)$ in the state $s$. In the following, we call the set of traits of an individual *the culture of an individual*.

**Example 1.** Let us consider a set of people and model them as agents with a set of traits and a behavior related to transmission, *telling_DA* (telling that Dante Alighieri wrote "The Divine Comedy"). Let $Ag$ in our example be a set of people: Charlie, Pedro, Maria, and Andrea are European citizens, and Toru is from Japan. Let $\mathcal{T}$ be a set of traits of different types.

Table 5 lists the sets of traits of the specific agents of $Ag = \{Charlie, Pedro, Toru, Maria, Andrea\}$ in the state $s_1$. Changes with respect to Table 2 are highlighted in bold. We can write $has(Maria, Dante\_Alighieri\_wrote\_Divine\_Comedy, s_1)$, or $has(Charlie, cappuccino\_is\_coffee, s_1)$, but not $has(Andrea, eating\_with\_sticks, s_1)$. We will use this example as a running example. ⊘

| set | traits |
| --- | --- |
| $T_{Charlie}(s_1)$ | *Dante_Alighieri_wrote_Divine_Comedy*, *latte_macchiato_is_coffee*, **telling_DA**, *cappuccino_is_coffee*, *eating_with_sticks*, *eating_with_fork*, *never_put_mayonnaise_on_pizza*, *Buddhism* |
| $T_{Pedro}(s_1)$ | *Dante_Alighieri_wrote_Divine_Comedy*, *latte_macchiato_is_coffee*, *cappuccino_is_coffee*, *eating_with_fork*, *Christianity* |
| $T_{Toru}(s_1)$ | *Meiji_era_was_in_1868_1912*, *cappuccino_is_coffee*, *eating_with_sticks*, *Buddhism* |
| $T_{Maria}(s_1)$ | *Dante_Alighieri_wrote_Divine_Comedy*, *latte_macchiato_is_coffee*, *cappuccino_is_coffee*, *eating_with_sticks*, *eating_with_fork*, *Christianity* |
| $T_{Andrea}(s_1)$ | *Dante_Alighieri_wrote_Divine_Comedy*, *latte_macchiato_is_coffee*, *cappuccino_is_coffee*, *eating_with_fork*, *Christianity* |

**Table 5.** Traits of agents in Example 1.

We distinguish behavior as a particular kind of traits and assume that performing a behavior by an agent changes the state of the world. We define the set of all behaviors $\mathcal{B} \subseteq \mathcal{T}$ and the function $perform$ in $Ag \times \mathcal{B} \times S \to S$. The intended meaning of this function is that an agent, which has some behavior in some state, performs this behavior in this state and the state of the world changes to another state. More specifically, $s_v = perform(a, \tau, s_u)$ means that $has(a, \tau, s_u)$ and the agent $a$ performed a behavior $\tau$ in the state $s_u$ and the resulting state is $s_v$. The fact that $has(a, \tau, s_u)$ does not imply that the agent $a$ is able to perform the behavior $\tau$ in the state $s_u$, because some preconditions for performing the behavior may be not fulfilled in the state $s_u$. Note that since traits are not innate, by assuming $\mathcal{B} \subseteq \mathcal{T}$ we do not include innate behaviors, such as blinking when air is puffed in someone's eye.

At this point we would like to discuss the distinction between *action* and *behavior*. In AI literature, an action is an atomic piece of activity, while behavior is perceived as something more complex, and can include several actions. Therefore, our notion of performing a behavior can really be decomposed into performing several actions. However, we decided not to introduce explicit relations between actions and behaviors. Moreover, the absence of such clear dependency in AI literature suggests that these relations are hard or even impossible to formalize. Instead, we assume that behavior can represent an atomic action or a more complex activity depending on the level of modeling granularity. We can vary granularity of behaviors depending on the problem in hand and on the domain. For instance, in Example 1, when someone needs to know whether agents are working, it is possible to consider behaviors *working* and *playing*, or, even, *working* and *not_working*. However, if someone would like to have a closer look at leisure activities of the group, it is necessary to introduce finer granularity of the *playing* behavior, e.g. by considering *playing_basketball* and *playing_chess* behaviors.

We assume that the states are ordered, we define recursively the order "is before" and the corresponding predicate $is\_before(s_u, s_v)$ and $is\_after(s_v, s_u)$ in the following way:

**Definition 6 (is_before).** $is\_before(s_u, s_v) \leftrightarrow \exists a \in Ag,\ \tau \in \mathcal{B},\ s \in S$ *such that* $s = perform(a, \tau, s_u) \wedge (s = s_v \vee is\_before(s, s_v))$.

**Definition 7 (is_after).** $is\_after(s_v, s_u) \leftrightarrow is\_before(s_u, s_v)$

We assume that in each state $s_v$, the previous state $s_u$ is uniquely defined, while the next state depends on the action an agent performs in $s_v$. From Definition 6 we can derive the following property:

*Property 4.* For all agents $a \in Ag$, for all behaviors $\tau \in \mathcal{B}$ and for all states $s_u, s_v \in S$

$$s_v = perform(a, \tau, s_u) \rightarrow is\_before(s_u, s_v)$$

**Definition 8 (sharing).** *For each pair of agents* $a_i, a_j \in Ag$*, for each trait* $\tau \in \mathcal{T}$*, and for each state* $s \in S$*,* $a_i$ *and* $a_j$ *share the trait* $\tau$ *in the state* $s$ *iff they both have such a trait in* $s$*:*

$$has(a_i, \tau, s) \wedge has(a_j, \tau, s) \leftrightarrow sharing(a_i, a_j, \tau, s).$$

We also assume that agents do not lose traits when the state of the world changes, as the following axiom says:

**Axiom 1** *For all agents* $a \in Ag$*, traits* $\tau \in \mathcal{T}$*, and states* $s \in S$ :

$$has(a, \tau, s) \rightarrow \forall s_v : is\_after(s_v, s)\ has(a, \tau, s_v).$$

**Example 1 (continued).** In the example, we can write $sharing(Toru, Maria, eating\_with\_sticks, s_1)$, or $sharing(Pedro, Andrea, cappuccino\_is\_coffee, s_1)$, etc. Traits that are shared by a pair of agents in the state $s_1$ are as in Figure 1. $\oslash$

Let us assume that if an agent $a_i$ has a trait $\tau$, the trait $\tau$ can be transmitted to another agent $a_j$ before some state $s$ and we use the predicate $transmitted(a_i, a_j, \tau, s)$ to represent this. We represent $transmitted(a_i, a_j, \tau, s)$ in a graph by a directed edge from $a_i$ to $a_j$ labeled $\tau$.

**Axiom 2** *For each pair of agents* $a_i, a_j \in Ag$*,* $a_i \neq a_j$*, for each trait* $\tau \in \mathcal{T}$*, and for each state* $s \in S$ *the fact that the trait* $\tau$ *has been transmitted from* $a_i$ *to* $a_j$ *before the state* $s$ *implies that exists some state* $s_u \in S$ *such that* $a_i$ *has* $\tau$ *in the state* $s_u$*,* $a_j$ *does not have* $\tau$ *in the state* $s_u$ *and an agent* $a_k$ *performing a behavior* $\tau_m$ *in the state* $s_u$ *imply that in the resulting state* $s_v$ *the agent* $a_j$ *has* $\tau$*:*

$$transmitted(a_i, a_j, \tau, s) \rightarrow (\exists s_u \in S,\ is\_before(s_u, s) \wedge has(a_i, \tau, s_u) \wedge \\ \neg has(a_j, \tau, s_u) \wedge (s_v = perform(a_k, \tau, s_u)) \rightarrow has(a_j, \tau, s_v))$$

From our assumption that traits are not innate, it follows that traits are acquired by agents, and the goal of the transmitted predicate is to show the way an agent acquired a trait. Therefore, we assume that in the initial state agents have no traits and the way they acquire traits is represented using the transmitted predicate.

We should note that the trait $\tau$ is not shared by $a_i$ and $a_j$ in the state $s_u$, while it is shared by $a_i$ and $a_j$ in the state $s_v$, and in the state $s$, as shown by the following property:

*Property 5.* For all pairs of agents $a_i, a_j \in Ag$, for all traits $\tau \in \mathcal{T}$, and for all states $s_v \in S$

$$sharing(a_i, a_j, \tau, s_v) \rightarrow (\forall s : is\_after(s, s_v)\ sharing(a_i, a_j, \tau, s))$$

From Axiom 2 it also follows that the transmitted predicate holds for all subsequent states after $s_v$.

*Property 6.* For all pairs of agents $a_i, a_j \in Ag$, for all traits $\tau \in \mathcal{T}$, and for all states $s_v \in S$

$$transmitted(a_i, a_j, \tau, s_v) \rightarrow (\forall s : is\_after(s, s_v)\ transmitted(a_i, a_j, \tau, s))$$

**Example 1 (continued).** Figure 3 shows the graph representing the *transmitted* predicate in state $s_1$ in our example. The traits $Dante\_Alighieri\_wrote\_Divine\_Comedy$ and $eating\_with\_sticks$ have been transmitted. On the contrary, the traits $cappuccino\_is\_coffee$ and $never\_put\_mayonnaise\_on\_pizza$ have not been transmitted (the latter trait is not even shared by any pair of agents). In particular, the $Dante\_Alighieri\_wrote\_Divine\_Comedy$ trait has been transmitted from $Charlie$ to $Maria$, and from $Maria$ to $Andrea$. Also, the $eating\_with\_sticks$ trait has been transmitted from $Charlie$ to $Toru$ and from $Toru$ to $Maria$. We can write $transmitted(Charlie, Maria, Dante\_Alighieri\_wrote\_Divine\_Comedy, s_1)$.

Let us assume that in the state $s_1$ $Charlie$ tells $Toru$ that Dante Alighieri wrote the Divine Comedy and $Toru$ memorizes this piece of knowledge. This corresponds to $s_2 = perform(Charlie, telling\_DA, s_1)$. The *transmitted* predicate in the state $s_2$ is as depicted in the left part of Figure 3 and *transmitted* in the state $s_2$ is as depicted in the right part of Figure 3. The difference in the *transmitted* predicates in these two states is that the $Dante\_Alighieri\_wrote\_Divine\_Comedy$ trait has been transmitted from $Charlie$ to $Toru$ and the corresponding edge is added, namely $transmitted(Charlie, Toru, Dante\_Alighieri\_wrote\_Divine\_Comedy, s_2)$. In the state $s_2$ the following change in the set of traits for Toru occurs: $T_{Toru}(s_2) = \{Meiji\_era\_was\_in\_1868\_1912, Dante\_Alighieri\_wrote\_Divine\_Comedy, cappuccino\_is\_coffee, eating\_with\_sticks, Buddhism\}$.

Obviously, the transmission has an impact on sharing and the *sharing* predicate in the state $s_2$ is as depicted in Figure 4, with the edges between $Toru$ and $Charlie$, $Maria$, $Andrea$, $Pedro$ added. ⊘

Given a set of agents $G \subseteq Ag$ and a set of traits $T_G \subseteq \mathcal{T}$ we define the notions of *weak sharing* and *strong sharing*.

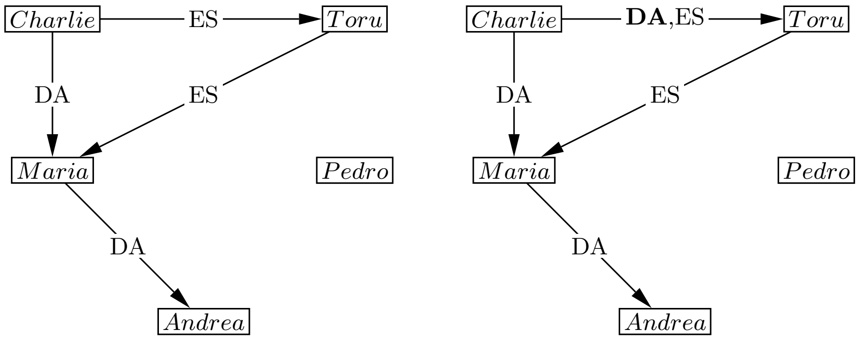


**Fig. 3.** The graph that shows for which agents the *transmitted* predicate holds in the state $s_1$ (left) and $s_2$ (right) in Example 1. Changes with respect to the state $s_1$ are in bold.

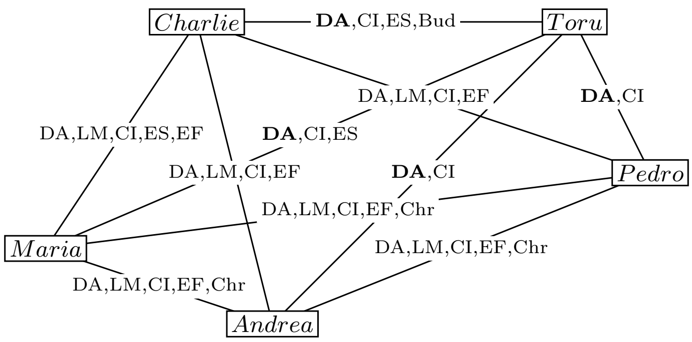


**Fig. 4.** The graph that shows for which agents the *sharing* predicate holds in the state $s_2$ in Example 1. Changes with respect to the state $s_1$ are in bold.

**Definition 9 (weak sharing).** *A set of traits $T_G$ is weakly shared by a set of agents $G$ in a state $s$ iff for each trait $\tau \in T_G$ there exists a pair of agents $a_i, a_j \in G$, $a_i \neq a_j$ that share $\tau$ in the state $s$.*

**Definition 10 (strong sharing).** *A set of traits $T_G$ is strongly shared by a set of agents $G$ in a state $s$ iff each trait $\tau \in T_G$ is shared by all pairs of agents $a_i, a_j \in G$ in $s$.*

In other words, the set of traits is weakly (strongly) shared if it is a subset of the union (intersection) of traits shared by pairs of agents of $G$ in the state $s$.

**Example 1 (continued).** Let us consider the set of agents $G = \{Charlie, Toru, Maria, Andrea, Pedro\}$. Analyzing the *sharing* predicate in the state $s_1$ we can see that only the $cappuccino\_is\_coffee$ trait is shared by each pair of agents in the state $s_1$, so $T_G = \{cappuccino\_is\_coffee\}$ is strongly shared by $G$ in the state $s_1$. There are three traits that are shared by at least one pair of agents in the state $s_1$: $cappuccino\_is\_coffee$, $eating\_with\_sticks$ shared, e.g, by $Toru$ and $Charlie$, and $Dante\_Alighieri\_wrote\_Divine\_Comedy$ shared, e.g., by $Charlie$ and $Andrea$. So, the set $T'_G = \{Dante\_Alighieri\_wrote\_Divine\_Comedy, cappuccino\_is\_coffee, eating\_with\_sticks\}$ and all non-empty subsets of this set are weakly shared by the set $G$ in the state $s_1$. Analogously, the set $T''_G = \{eating\_with\_sticks, Dante\_Alighieri\_wrote\_Divine\_Comedy, cappuccino\_is\_coffee\}$ is weakly shared by $G$ in the state $s_3$, and the set $T'''_G = \{cappuccino\_is\_coffee, Dante\_Alighieri\_wrote\_Divine\_Comedy\}$ is strongly shared by the set $G$ in the state $s_3$. ⊘

*Property 7.* Strong sharing implies weak sharing.

Given a set of agents $G \subseteq Ag$ such that $|G| \geq 2$, and a *transmitted* predicate we introduce the notion of culture of $G$.

**Definition 11 (weak culture of a set of agents).** *A non-empty set of traits $T_G \subseteq \mathcal{T}$ is a weak culture of $G$ in a state $s$ iff*

- *the set $T_G$ is weakly shared by $G$ in the state $s$,*
- *for each agent $a \in G$ in the state $s$ there exists a trait $\tau \in T_G$ such that $has(a, \tau, s)$.*

From the assumption that traits are not innate, as we discussed, it follows that traits are acquired by agents, as represented by the transmitted predicate. Therefore, we can formulate the following axiom, telling that all traits in culture are transmitted.

**Axiom 3** *For each trait $\tau \in T_G$ there exists an agent $a \in Ag$ that transmitted $\tau$ to another agent $a_j \in G$ before the state $s$, i.e. $transmitted(a, a_j, \tau, s)$.*

From Definition 11 and Axiom 3 it follows that all the traits in the culture are transmitted, shared, and each agent has at least one trait from the culture. Please, note that since the traits are transmitted not necessarily within the set, the transmitted predicate does not imply sharing between the agents of $G$.

**Definition 12 (strong culture of a set of agents).** *If $T_G$ in Definition 11 is also strongly shared in the state $s$ then it is* a strong culture of the set of agents *$G$ in the state $s$.*

In the following if we refer to "a culture of a set of agents", we mean "a weak culture of a set of agents".

**Example 1 (continued).** Considering $G = \{Toru, Andrea\}$ in the state $s_3$, $T_G = \{Dante\_Alighieri\_wrote\_Divine\_Comedy,\ cappuccino\_is\_coffee\}$ is strongly shared by the set $G$ in the state $s_3$.

Although the $Dante\_Alighieri\_wrote\_Divine\_Comedy$ trait has been transmitted both to $Toru$ and $Andrea$ from outside (from $Charlie$ and $Maria$, respectively), it is strongly shared by the agents of $G$. Since in the state $s_3$ each agent in $G$ has the trait $Dante\_Alighieri\_wrote\_Divine\_Comedy$, $T'_G = \{Dante\_Alighieri\_wrote\_Divine\_Comedy\}$ is a culture of $G$ in the state $s_3$. It is easy to see that $T'_G$ is not a culture of $G$ in the states $s_1$ and $s_2$ because $Toru$ does not have this trait in those states. ⊘

The following proposition outlines some restrictions on how culture can change between states, namely it shows that culture is monotonic.

**Proposition 1 (monotonicity of culture).** *If a non-empty set of traits $T_G$ is a culture of a set of agents $G$ in a state $s_v$, then $T_G$ is a culture of $G$ also in any state $s$ after $s_v$.*

In real world, the traits of a culture can be lost for two reasons: (1) agents can lose traits, (2) agents can die, move to another group, etc. As we stated in Axiom 1, in our model, agents do not lose traits. However, our model, and the proposition about monotonicity of culture support the case when agents disappear from the group.

**Definition 13 (union culture of a group).** *A non-empty set of traits $T_G^{union}$ is the union culture of a set of agents $G$ in the state $s$ iff $T_G^{union}$ is the union of all cultures $T_G$ of $G$ in the state $s$.*

In other words, the union culture of a set of agents in some state is the union of all possible cultures of the set in this state. Since it is the union of all cultures, it is not possible to add any trait to $T_G^{union}$ and still obtain a culture of $G$. In the following, we refer to the union culture of a set of agents as "the culture of a set".

**Definition 14 (evolution of culture).** *A sequence of sets of traits $\{T_G^{(1)}, \ldots, T_G^{(i)}\}$ is an evolution of culture of $G$ iff:*

- *exists a sequence of states $\{s_1, \ldots, s_i\}$, such that $T_G^{(k)}$ is a culture of $G$ in the state $s_k$ for all $k$, $1 \leq k \leq i$,*
- *for each $k$, $1 \leq k \leq i-1$ holds $is\_after(s_{k+1}, s_k)$.*

In other words, a sequence of sets of traits is an evolution of culture if each set of traits in the sequence is a culture of $G$ in some state and the states are ordered in the same way as the sets of traits. We denote evolution of culture as $\{T_G\}$.

## 7 Related work and discussion

There are a number of approaches that are related to the development of applications for communities: social navigation [22], community-based personalization [23], open source software development to mention a few. However, existing approaches develop applications that support the culture pre-defined by the designer rather than actual culture of the community where the application is being deployed. For instance, in 1998 NetCaptor browser, followed by Opera in 2000 and Mozilla Firefox in 2001, introduced tabbed browsing, thus opening a new way to browsing in the Internet. As time passed, more and more people became familiar with tabbed browsing and this lead other browsers, e.g. IE Explorer, to introduce tabbed browsing. We argue that tabbed browsing can be considered an element of culture, rarely found in the culture of Internet users of 90s, but very common nowadays[3], and this change forced Microsoft to change the pre-defined culture of supporting only one-page-one-window browsing to the actual culture of tabbed browsing.

The AI literature on agents addresses the issues related to sociality, such as social action [24], social co-ordination architectures and social interaction strategies for decentralized co-ordination in multi-agent systems [25], social laws and conventions in multi-agent environments [26, 27], and social roles [28]. However, the issue of sociality alone does not help neither to understand what differentiates one set of agents from another nor to grasp what are the specificities of the behavior of agents of a specific society. Although in two different agent societies agents can be able to communicate with each other and perform other social actions, these two societies can be very different from each other. We claim that the concept of culture can be used to describe and compare sets of agents. Some research papers use the notion of culture in the context of agents, see e.g. [1, 21, 29], other provide a model for the comparison of cultures [30]. However, none of the previous research works provides a formal definition of culture that could be readily adopted for building applications for communities and applied for the characterization and the comparison of culture.

Carley [21] considers culture as the distribution of information (ideas, beliefs, concepts, symbols, technical knowledge, etc.) across the population and proposes a model for knowledge transfer based on interactions. In that model, the probability of an interaction between two agents is based on the principle of homophily, i.e. the greater the amount of knowledge they share the more probable the interaction is. During an interaction, agents exchange facts, so after the interaction one of the agents might know more than before the interaction. The knowledge transfer in these settings can be seen as a particular kind of culture spread. This work is further extended in the Construct project [31, 32]. For instance, one of the recent applications of Construct studies the effects of different methods of information diffusion on spreading beliefs and knowledge about illegal tax schemes in different American cities [33]. With respect to the definition of culture we propose in this paper, that model of information diffusion is com-

[3] For instance, Google Chrome browser includes tabbed browsing from the first release

plementary, because it models transmission of elements of culture (e.g., beliefs, knowledge) in a society.

Axelrod [1] considers culture as a list of features or dimensions of culture. Each feature represents an individual attribute that is subject to social influence and can have different values called traits. Two individuals have the same culture if they have the same traits for all features. Similarly to the work by Carley, feature of an agent can change its value during an interaction and the probability of interaction is based on the homophily.

The notion of trait we use in our formalism is similar to the notion of feature used by Axelrod, specifically, each feature can take value from a set of specific traits. Traits in our formalism also includes ideas, beliefs and technical knowledge used as culture elements by Carley. Both theories by Carley and by Axelrod are based on the assumption that culture changes as a result of an interaction. Thus, in our terms, interaction in that sense can be considered as a particular kind of transmission: there are two agents participating, it takes place in some specific state and it leads to the appearance of some cultural element in one of the agents.

Epstein and Axtell [29] study the emergence of the group rules from local ones defined at an agent's level in an artificial society of simple agents living and consuming sugar in an artificial environment called "Sugarscape". The authors consider a culture of the society as a string of binary cultural attributes and model cultural transmission both on horizontal (between agents) and vertical (through generations) levels using simple rules. However, they do not provide any formal definition of culture since the main focus of the book is on the emergence of group rules from the local ones.

According to O'Reilly [7], the culture of an organization is considered as strong if wide consensus exists about the content and participants believe in the importance of the content. They also formulate this as a [not necessarily big] set of values that are widely shared and strongly held. This is similar to the notion of strong culture, i.e. culture shared by all pairs of agents in a group, we consider in our formalism.

Balzer and Tuomela [34] study social practices and the dynamics of their maintenance in groups. They define social practices as recurrent collective activities based on collective intentions. The paper focuses on informal, non-normative practices, such as playing soccer on Sundays, going to sauna on Saturday afternoon, shaking hands, sharing a ride to work. They also note that the maintenance (change, preservation, renewal) depend on the success of a practice. The main contribution of the paper is a mathematical model for the description of social practices and their maintenance in groups.

Our model of culture is not limited to social practices. Moreover, it allows for inclusion of normative practices as well. However, as a consequence, the model of Balzer and Tuomela allows for a richer description of informal social practices. For instance, our model does not permit expressing intentions, but allows operating on manifestations of activities without going into details of underlying intentions. While authors show that success of a social practice is important for its adoption, for our model it is irrelevant whether a trait is successful in some sense. Our model just captures the fact that the trait is a part of culture, no matter how it occurred. The model presented by Balzer and Tuomela is defined for groups and then goes to the individual level, thereby implementing top-down approach. In our model of culture, we start from a set of traits of an individual, consider transmission as an important means of spreading culture, and then go to the culture of a group. Thus, we implement bottom-up approach. Balzer and Tuomela, while requiring sharing of a social practice within a group, and noting the importance of transmission for spreading practice, include transmission into the model only to a certain extent, namely, considering imitation as an example of transmission. Our model of culture allows for different types of transmission as long as there is a predicate that helps to distinguish occurred transmissions.

Hofstede [30] treats culture as "[...] the collective programming of the mind that distinguishes the members of one group or category of people from another", proposes a model of culture and applies it for studying and comparing cultures of IBM workers in more than 50 countries. The model includes the following five independent dimensions of national culture differences: *power distance*, which is related to the different solutions to the basic problem of human inequality; *unvertainty avoidance*, which is related to the level of stress in a society in the face of an unknown future; *individualism* versus *collectivism*, which is related to the integration of individuals into primary groups; *masculinity* versus *femininity*, which is related to the division of emotional roles between men and women; and *long-term* versus *short-term* orientation, which is related to the choice of focus for people's efforts: the future or the present. Values in Hofstede's terms refer to "a broad tendency to prefer certain states over others" and are similar to attitudes and beliefs, which are just particular kind of traits in our formalism. Dimensions, similarly to Axelrod's features, take values from the set of traits. Thus, comparing with our work, the model developed by Hofstede has a different focus - it aims at comparing cultures of groups of people over several pre-defined dimensions of values, while our model supports comparison over arbitrary sets of traits. The dimensions in Hofstede's model are meant to be independent, while our formalism does not address the issue of dependency of traits, so they can be dependent on each other. In this line of thoughts, an interesting application of our model could be comparison of dependency of traits across groups, i.e. if presence of a trait or traits leads to the presence of another trait(s) for one group and to the presence of third trait(s) for another group.

The definition of culture presented here allows for representation and comparison of different cultures. However, in order to compare traits, one first needs to identify the traits of individuals. On the one hand, deducing traits from manifested behaviors of agents is not a trivial task in general. On the other hand, in specific domains this might be much easier, consider, for instance, deducing traits of users from logs of a web service, website, or an application. For instance, it would be possible to see that a group of users of a text editor always turn off the autocorrect feature and turn it off automatically in new versions of the editor prepared for this group. Taking the issue of the observability of traits into account, we see social software and Web 2.0 systems as one of the potential application domains for our model.

## 8 Conclusions

We have defined the notion of culture of a set of agents and we have shown that our definition can be used for comparison of communities. This definition addresses existing gaps in AI literature that deals with issues of sociality, cooperation, and negotiation, but remains oblivious to the notion of culture. The provided formalism is a first step towards an integral approach for representing, comparing, analysing, and transferring culture of communities or group of agents. We are currently working on the extension of the formalism with the notion of culture evolution to study culture changes in Web 2.0 communities.


## Acknowledgments

We would like to thank Stefano Borgo for suggestions he provided for improving the formalism.


## References


1. Axelrod, R.: The dissemination of culture: A model with local convergence and global polarization. The Journal of Conflict Resolution **41**(2) (1997) 203–226
2. Mulder, M.B., Nunn, C.L., Towner, M.C.: Cultural macroevolution and the transmission of traits. Evolutionary Anthropology **15**(2) (2006) 52–64
3. Birukou, A., Blanzieri, E., Giorgini, P., Giunchiglia, F.: A formal definition of culture of a set of agents. the version without states. Technical Report DISI-09-070, University of Trento (2009)
4. Birukou, A., Blanzieri, E., Giorgini, P., Giunchiglia, F.: A formal definition of culture. In: Proceedings of the Workshop on Modeling Intercultural Collaboration and Negotiation (MICON) at IJCAI'09. (2009)
5. Mead: Cooperation and Competition Among Primitive Peoples. Transaction Publishers (2002)
6. Brumann, C.: Writing for culture: Why a successful concept should not be discarded [and comments and reply]. Current Anthropology **40** (1999)
7. O'Reilly, C.: Corporations, culture, and commitment: Motivation and social control in organizations. California Management Review **31**(4) (1989) 9–25
8. Mead, M., Métraux, R., eds.: The Study of Culture at a Distance. University of Chicago Press (1953)
9. McFate, M.: The military utility of understanding adversary culture. Joint Force Quarterly **38** (2005)
10. Bozeman, A.B.: Strategic Intelligence and Statecraft: Selected Essays. Washington, DC: Brassey's (1992)
11. Boyd, R., Richerson, P.J.: Culture and the Evolutionary Process. University of Chicago Press (1988)

12. Kroeber, A.L., Kluckhohn, C.: Culture: A critical review of concepts and definitions. Papers of the Peabody Museum **47** (1952) 643–656
13. Harrison, R.J., Carroll, G.R.: Culture and Demography in Organizations. Princeton University Press (2006)
14. Schein, E.H.: Organizational Culture and Leadership (Jossey-Bass Psychology Series). 2 edn. Jossey-Bass (1996)
15. Bailey, G., Peoples, J.: Introduction to Cultural Anthropology. Wadsworth Publishing (1998)
16. Keesing, R.M.: Cultural anthropology: A contemporary perspective. Holt, Rinehart and Winston (1981)
17. Tylor, E.B.: Primitive culture. London: J. Murray (1871)
18. Linton, R.: The study of man: An introduction. New York: D. Appleton-Century (1936)
19. Harris, M.: Culture, people, nature: An introduction to general anthropology. New York: Thomas Y. Crowell (1975)
20. Ferber, J.: Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence. Addison-Wesley Professional (1999)
21. Carley, K.: A theory of group stability. American Sociological Review **56**(3) (1991)
22. Dieberger, A.: Supporting social navigation on the world wide web. Int. J. Human-Computer Studies **46** (1997) 805–825
23. Freyne, J., Smyth, B., Coyle, M., Balfe, E., Briggs, P.: Further experiments on collaborative ranking in community-based web search. Artificial Intelligence Review **21**(3-4) (2004) 229–252
24. Castelfranchi, C.: Modelling social action for ai agents. Artificial Intelligence **103**(1-2) (1998) 157–182
25. Ossowski, S.: Co-ordination in Artificial Agent Societies, Social Structure and Its Implications for Autonomous Problem-Solving Agents. Volume 1535 of Lecture Notes in Computer Science. Springer (1999)
26. Shoham, Y., Tennenholtz, M.: On social laws for artificial agent societies: off-line design. Artif. Intell. **73**(1-2) (1995) 231–252
27. Shoham, Y., Tennenholtz, M.: On the emergence of social conventions: modeling, analysis, and simulations. Artificial Intelligence **94**(1-2) (1997) 139–166
28. Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., Guarino, N.: Social roles and their descriptions. In: Principles of Knowledge Representation and Reasoning (KR2004). (2004)
29. Epstein, J.M., Axtell, R.L.: Growing Artificial Societies: Social Science from the Bottom Up. The MIT Press (1996)
30. Hofstede, D.G.: Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations. 2nd edn. Sage Publications, Inc (April 2001)
31. Hirshman, B.R., Carley, K.M., Kowalchuck, M.J.: Specifying agents in construct. Technical Report CMU-ISRI-07-107, Carnegie Mellon University (2007)
32. Hirshman, B.R., Carley, K.M., Kowalchuck, M.J.: Loading networks in construct. Technical Report CMU-ISRI-07-116, Carnegie Mellon University (2007)
33. Hirshman, B.R., Birukou, A., Martin, M.K., Bigrigg, M.W., Carley, K.M.: The impact of educational interventions on real & stylized cities. Technical Report CMU-ISR-08-114, Carnegie Mellon University (2008)
34. Balzer, W., Tuomela, R.: Collective intentions and the maintenance of social practices. Autonomous Agents and Multi-Agent Systems **6**(1) (2003) 7–33