Department of
Information Engineering
and Computer Science DISI
DISI - Via Sommarive, 14 - 38123 POVO, Trento - Italy
http://disi.unitn.it

UNIVERSITY OF TRENTO - Italy

A FORMAL DEFINITION OF CULTURE OF A SET OF AGENTS. THE VERSION WITHOUT STATES.

Aliaksandr Birukou, Enrico Blanzieri, Paolo Giorgini, Fausto Giunchiglia

December 2009

Technical Report # DISI-09-070

A formal definition of culture of a set of agents. The version without states.

Aliaksandr Birukou [*], Enrico Blanzieri, Paolo Giorgini, Fausto Giunchiglia

January 14, 2010

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. Most current approaches to the development of systems for online communities implement some pre-defined culture rather than supporting the actual culture of the community for which the system is being developed.

This technical report presents a formal definition of culture of a set of agents. Our definition covers many previous definitions existing in anthropology, and it can be used to model the culture of online communities. This is the first version of the formalism and it 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).

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

[*]If you are interested in the work, or would like to comment or collaborate, please do not hesitate to contact the first author

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 and occur 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 technical report 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 is the first version of the formalism and it does not introduce states and does not consider spreading of culture. However, representing a snapshot of the culture in a certain moment is the first step towards a more complex formalism that includes states and models transmission. Therefore, this technical report is a precursor to the version of the formalism with the states, presented in [3] (see [4] for a slightly updated version). Also, please refer to those papers for a discussion of relations between our formalism and existing anthropological and computer science research.

In Section 2 we provide a formal definition of the culture of a set of agents and of the related concepts. Section 3 defines some measures on top of the formalism, and an example, illustrating our approach, is presented in Section 4. We describe related work in Section 5, and conclude the report in Section 6.

2 A formal definition of culture

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 tenden-

trait type	traits
knowledge	$Dante\_Alighieri\_wrote\_Divine\_Comedy(DA)$, $latte\_macchiato\_is\_coffee(LM)$, $Meiji\_era\_was\_in\_1868\_1912(ME)$, $cappuccino\_is\_coffee(CI)$
behavior	$eating\_with\_sticks(ES)$, $eating\_with\_fork(EF)$, $taking\_vacation\_in\_August(TVA)$, $taking\_vacation\_in\_May(TVM)$
norms, rules	$never\_put\_mayonnaise\_on\_pizza(NP)$, $take\_only\_week\_of\_vacation\_per\_year(T1W)$, $never\_drink\_cappuccino\_after\_lunch(ND)$, $never\_open\_umbrella\_inside\_building(NO)$
beliefs	$Christianity(Chr)$, $Buddhism(Bud)$

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

cies" [5]. 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*.

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 $\oslash$

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 *taking_vacation_in_August* 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.

set	traits
$\mathcal{T}$	$Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $Meiji\_era\_was\_in\_1868\_1912$, $cappuccino\_is\_coffee$, $eating\_with\_sticks$, $eating\_with\_fork$, $taking\_vacation\_in\_August$, $taking\_vacation\_in\_May$, $never\_put\_mayonnaise\_on\_pizza$, $take\_only\_week\_of\_vacation\_per\_year$, $never\_drink\_cappuccino\_after\_lunch$, $never\_open\_umbrella\_inside\_building$, $Christianity$, $Buddhism$
$T_{Charlie}$	$Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $cappuccino\_is\_coffee$, $eating\_with\_sticks$, $eating\_with\_fork$, $taking\_vacation\_in\_August$, $never\_put\_mayonnaise\_on\_pizza$, $Buddhism$
T_{Pedro}	$Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $cappuccino\_is\_coffee$, $eating\_with\_fork$, $taking\_vacation\_in\_August$, $never\_drink\_cappuccino\_after\_lunch$, $Christianity$
T_{Toru}	$Meiji\_era\_was\_in\_1868\_1912$, $cappuccino\_is\_coffee$, $eating\_with\_sticks$, $taking\_vacation\_in\_May$, $Buddhism$
T_{Maria}	$Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $cappuccino\_is\_coffee$, $eating\_with\_sticks$, $eating\_with\_fork$, $taking\_vacation\_in\_August$, $Christianity$
T_{Andrea}	$Dante\_Alighieri\_wrote\_Divine\_Comedy$, $latte\_macchiato\_is\_coffee$, $cappuccino\_is\_coffee$, $eating\_with\_fork$, $taking\_vacation\_in\_August$, $Christianity$

Table 2: Traits of agents in Example 1.

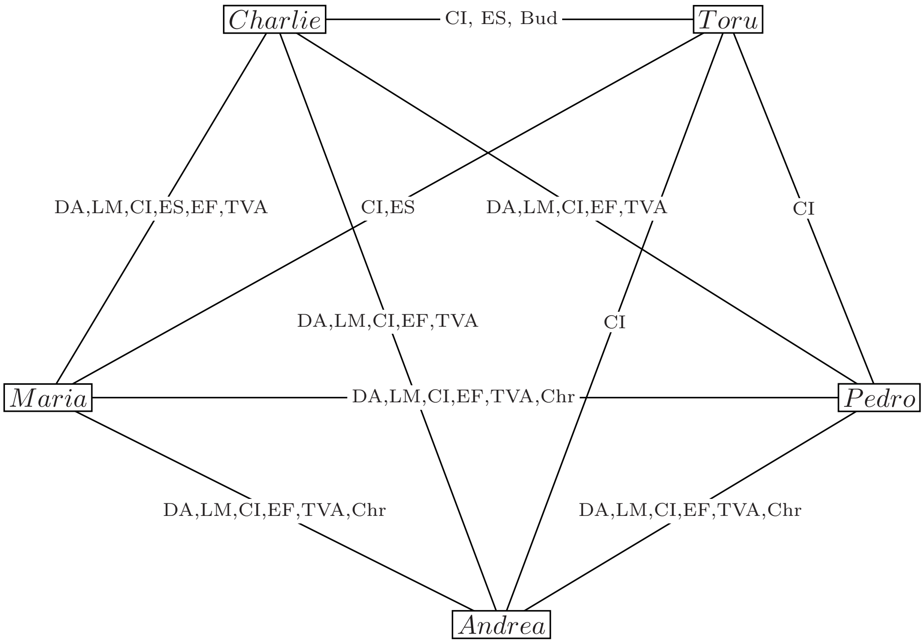

Figure 1: The graph showing for which agents and traits the predicate *sharing* holds in Example 1. The nodes are agents and labels on each edge denote traits that are shared by the pair of agents connected by the edge. For instance, the edge between $Toru$ and $Andrea$ labeled CI means that $sharing(Andrea, Toru, cappuccino\_is\_coffee)$. The traits are abbreviated as in Table 1, i.e., *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, *taking_vacation_in_August* as TVA, *Christianity* as Chr, *Buddhism* as Bud.

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 τ 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.

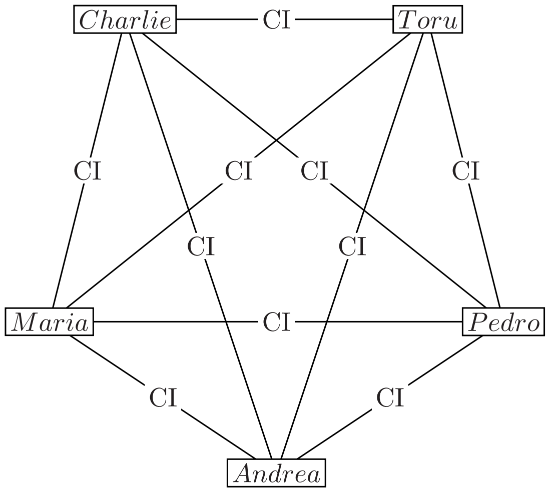

Figure 2: The graph that shows for which agents the *sharing* predicate holds for the $cappuccino\_is\_coffee(CI)$ trait in Example 1.

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. $\oslash$

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 τ.*

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 two sets of traits T_G =$\{cappuccino\_is\_coffee$, $eating\_with\_sticks$,$Dante\_Alighieri\_wrote\_Divine\_Comedy\}$, $T'_G = \{cappuccino\_is\_coffee\}$, and the set $G = \{Charlie, Toru, Maria, Andrea, Pedro\}$. Using the *sharing* predicate represented in Figure 1, we can see that the $cappuccino\_is\_coffee$ trait is shared by each pair of agents, so T'_G is strongly shared by G. 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. $\oslash$

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 τ. 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 τ is fulfilled. $\square$

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 $Dante\_Alighieri\_wrote\_Divine\_Comedy$ and $eating\_with\_sticks$ 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. $\oslash$

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. $\square$

3 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.

3.1 Measuring culture as a snapshot

3.1.1 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 τ 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.

3.1.2 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 [6]. 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 τ 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

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

In this formula, $G = \{a_i\},\ 1 \leq i \leq N$, $T_G = \{\tau_k\},\ 1 \leq k \leq K$, 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

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

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

3.1.3 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 τ_1, τ_2 we know the value of $sim(\tau_1, \tau_2)$. Culture similarity can then be defined as $CC(T_a, T_G) = \frac{\sum\limits_{i=1}^{|T_a|} \sum\limits_{j=1}^{|T_G|} sim(\tau_i^a, \tau_j^G)}{|T_a| \times |T_G|} \times 100\%$
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 $telling\_DA$.

measure	meaning
$I_{has}(\tau, T)$	shows if the trait τ 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.

3.1.4 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.

3.2 Example

Let us see how the described measures apply to Example 1 from Section 2, 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:

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

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

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

Proceeding with calculations we get:

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

4 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.

4.1 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[1]. He discovers that there are groups on CiteULike and that there are at

[1]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.

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 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.

4.2 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 3, 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.

4.3 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.

5 Related work

There are a number of approaches that are related to the development of applications for communities: social navigation [7], community-based personalization [8], 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[2], 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 [9], social co-ordination architectures and social interaction strategies for decentralized co-ordination in multi-agent systems [10], social laws in multi-agent environments [11], and social roles [12]. 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. [6, 13, 1], 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 [6] 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

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

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 [14, 15]. 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 [16]. With respect to our definition of culture, this model of information diffusion is complementary, 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 and also includes ideas, beliefs and technical knowledge used as culture elements by Carley.

Epstein and Axtell [13] 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 [17], 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.

Hofstede [18] 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; *uncertainty 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. 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.

6 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. 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 notions of states and culture evolution.

References

[1] R. Axelrod, The dissemination of culture: A model with local convergence and global polarization, The Journal of Conflict Resolution 41 (2) (1997) 203–226.
URL `http://links.jstor.org/sici?sici=0022-0027\%28199704\%2941\%3A2\%3C203\%3ATDOCAM\%3E2.0.CO\%3B2-A`

[2] M. B. Mulder, C. L. Nunn, M. C. Towner, Cultural macroevolution and the transmission of traits, Evolutionary Anthropology: Issues, News, and Reviews 15 (2) (2006) 52–64.
URL `http://dx.doi.org/10.1002/evan.20088`

[3] A. Birukou, E. Blanzieri, P. Giorgini, F. Giunchiglia, A formal definition of culture, in: Proceedings of the Workshop on Modeling Intercultural Collaboration and Negotiation (MICON) at IJCAI'09, 2009.

[4] A. Birukou, E. Blanzieri, P. Giorgini, F. Giunchiglia, A formal definition of culture, Tech. Rep. DISI-09-021, University of Trento (June 2009).
URL `http://eprints.biblio.unitn.it/archive/00001604/`

[5] J. Ferber, Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence, Addison-Wesley Professional, 1999.
URL `http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20\&path=ASIN/0201360489`

[6] K. Carley, A theory of group stability, American Sociological Review 56 (3) (1991) 331–354.
URL http://links.jstor.org/sici?sici=0003-1224\%28199106\%2956\%3A3\%3C331\%3AATOGS\%3E2.0.CO\%3B2-A

[7] A. Dieberger, P. Dourish, K. Hook, P. Resnick, A. Wexelblat, Social navigation: Techniques for building more usable systems, interactions 7 (6) (2000) 36–45.

[8] J. Freyne, B. Smyth, M. Coyle, E. Balfe, P. Briggs, Further experiments on collaborative ranking in community-based web search, Artificial Intelligence Review 21 (3-4) (2004) 229–252.

[9] C. Castelfranchi, Modelling social action for ai agents, Artificial Intelligence 103 (1-2) (1998) 157–182.
URL http://portal.acm.org/citation.cfm?id=1124699

[10] S. Ossowski, Co-ordination in Artificial Agent Societies, Social Structure and Its Implications for Autonomous Problem-Solving Agents, Vol. 1535 of Lecture Notes in Computer Science, Springer, 1999.
URL http://dblp.uni-trier.de/rec/bibtex/books/sp/Ossowski99

[11] Y. Shoham, M. Tennenholtz, On social laws for artificial agent societies: off-line design, Artif. Intell. 73 (1-2) (1995) 231–252.
URL http://portal.acm.org/citation.cfm?id=201244.201260

[12] C. Masolo, L. Vieu, E. Bottazzi, C. Catenacci, R. Ferrario, A. Gangemi, N. Guarino, Social roles and their descriptions, 2004.

[13] J. M. Epstein, R. L. Axtell, Growing Artificial Societies: Social Science from the Bottom Up, The MIT Press, 1996.
URL http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20\&path=ASIN/0262550253

[14] B. R. Hirshman, K. M. Carley, M. J. Kowalchuck, Specifying agents in construct, Tech. Rep. CMU-ISRI-07-107, Carnegie Mellon University (July 2007).

[15] B. R. Hirshman, K. M. Carley, M. J. Kowalchuck, Loading networks in construct, Tech. Rep. CMU-ISRI-07-116, Carnegie Mellon University (July 2007).

[16] B. R. Hirshman, A. Birukou, M. K. Martin, M. W. Bigrigg, K. M. Carley, The impact of educational interventions on real & stylized cities, Tech. Rep. CMU-ISR-08-114, Carnegie Mellon University (2008).

[17] C. O'Reilly, Corporations, culture, and commitment: Motivation and social control in organizations, California Management Review 31 (4) (1989) 9–25.
URL http://harvardbusinessonline.hbsp.harvard.edu/b01/en/common/item_detail.jhtml?id=CMR021

[18] D. G. Hofstede, Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations, 2nd Edition, Sage Publications, Inc, 2001.