IMPLEMENTING PERSONALITY ADJECTIVES AS BEHAVIORAL SCHEMES

François Bouchet and Jean-Paul Sansonnet
LIMSI-CNRS, BP 133 F-91403 Orsay cedex, France

ABSTRACT

In this paper we present a study dedicated to the computational implementation of personality traits in Conversational Agents. First, a significant set of personality-traits adjectives has been collected from thesaurus sources. Then the lexical semantics related to personality-traits has been extracted using the WordNet database, providing a formal representation in terms of so-called Behavioral Schemes.

KEYWORDS

Conversational agents, Personality-traits adjectives, Behavioral schemes.

1. INTRODUCTION

This study considers the particular context of conversational situations where three entities are in bilateral interaction: a human user (U), an assistant agent (A) and a computer system (S). Stemming from (Maes 1994) and (Cassell 2000), UAS situations cover a large class of conversational agents: Presenters, Helpers, Butlers, Friends, Companions, Teachers, Trainers and Coaches. A conversational agent, in the UAS situation, has two faces needing distinct capabilities: the *control* face, directed towards the system, and the *dialogical* face, directed towards the user.

— Controlling a computer application requires both a symbolic model of the application and a rational reasoning capacity over that model. In the following, we will refer to the control face of an agent as the ***rational agent*** in a way compliant with research in Artificial Intelligence (Russell 2009; Sabouret 2001).

— Dialoguing with the user requires three main elements: a) a conversational interface (often multimodal); b) the input of user's requests and the output of factual replies processed by the rational reasoning capacity; c) the expression of the agent's personality according to its actual role in one of the UAS situations listed above. In the following, we will refer to the expression of the agent's personality as the ***behavioral agent***.

Providing a conversational agent with a computational implementation of a personality model requires first to elicit the basic constituents composing a personality and second to define a framework for the computational implementation of those basic constituents.

— Traditionally, psychological studies have focused on three main classes of phenomena: a) low level *emotions* (*e.g.* emotions from (Ekman et al. 1972)); b) more stable notions related to *personality traits* (ranging from earlier research on personality adjectives/nouns by (Baumgarten 1933), (Anderson 1968) and (Allport & Odbert 1936) to the more recent classifications like OCEAN traits in the Five Factors Model (Goldberg 1981) (McCrae & Costa 1987); c) *static roles* involving normative social relationships and obligations (*e.g.* the user/agent roles in the UAS situation). In this study, we will focus on traits because they are intrinsic and stable, at least over a conversational session. While OCEAN provides a popular approach to trait classification, this model is far too general to exhibit the basic sought out constituents, so we rely on data offered by linguistic resources through the lexical semantics of thesaurus and databases such as WordNet (Fellbaum 1998), in a similar way to (Valitutti &Strapparava 2004-2008) works on 'Wordnet affect' or in our previous works on textual corpora for assistant agents (Bouchet 2007).

— In order to associate a computational implementation of the basic constituents of the personality traits we have to define a formalism to express the lexical semantics extracted from the linguistic resources. This is

achieved with the description of so-called *Behavioral Schemes*. Given an agent A with a set of capacities (*i.e.* a set $\mathbb{A}$ of atomic actions $\alpha_i \in \mathbb{A}$, that the agent can perform upon the system), a Behavioral Scheme $\sigma \in \mathbb{S}$ is defined as a symbolic representation of the psychological attitude of the agent with regard to its capacities.

In section 2 we present the process of elicitation of the lexical semantics associated with a class of personality-trait adjectives retrieved from thesaurus sources. In section 3 we describe the formal representation of the lexical semantics and we give some examples of personality-traits implementation.

2. PERSONALITY ADJECTIVES

The process of gathering a set of personality adjectives $p_i \in \mathbb{P}$ has been carried out to determine the most popular p_i in order to exhibit the most socio-cognitively salient σ_i. As a second order requirement, we also wanted a significant (but not necessarily exhaustive) coverage of p_i. This is the reason why, to collect the so-called Cp corpus, we relied on ten distinct Internet sources that explicitly claim to provide "lists of adjectives describing personality traits" (the sources are summarized in Table 1).

Table 1. Sources of personality adjectives. List 3 is quite large because we extended the coverage of Cp with one list of more general adjectives; List 7 is a merged list of two close-related sources.

Source	Word count
1. http://personal.georgiasouthern.edu/~jbjoy/Adjectives.html	315
2. http://www.keepandshare.com/doc/view.php?u=12894	182
3. http://www.esldesk.com/vocabulary/adjectives.htm (general adjectives)	733
4. http://www.lingolex.com/personalidad.htm	52
5. http://www.lesn.appstate.edu/fryeem/RE4030/character_trait_descriptive_adje.htm	183
6. http://www.mckinnonsc.vic.edu.au/la/lote/german/materials/describe/pers-adj.htm#top	363
7. http://www.learnenglish.de/grammar/adjectivepersonality.htm#positive merged with: http://www.examples-help.org.uk/parts-of-speech/personality-adjectives.htm	277
8. http://jobmob.co.il/blog/positive-personality-adjectives/ (biased by 'positiveness')	130
9. http://www.nonstopenglish.com/exercise.asp?exid=440	20
10. http://www.scribd.com/doc/2212798/Adjective-List	80
TOTAL	**2335**
TOTAL of union (lemma different)	**1303**

We have identified the most significant p_i by sorting the corpus according to the frequency of appearance of each p_i in the selected sources (hence the theoretical maximum score is 10 but the maximum observed one is only 9). In the resource files, adjectives are given as lemmas. Generally, to a lemma are associated several lexical semantics senses/meanings (noted /thesense/). The first way to assess the various meanings of a lemma is to consult a dictionary like the online version of the Merriam-Webster dictionary (MW). For example, MW gives the following description of the lemma 'friendly' with four main senses (the first one having three variants):

friend·ly \'fren(d)-lē\ Function: *adjective* Inflected Form(s): friend·li·er; friend·li·est Date: before 12th century

1 : of, relating to, or befitting a friend: as **a:** showing kindly interest and goodwill **b:** not hostile <a friendly merger offer>; *also* **:** involving or coming from actions of one's own forces <friendly fire> **c:** **CHEERFUL**, **COMFORTING** <the friendly glow of the fire>

2 : serving a beneficial or helpful purpose

3 : easy to use or understand <friendly computer software> — often used in combination <a reader-*friendly* layout>

4 : **COMPATIBLE**, **ACCOMMODATING** <environmentally friendly packaging> — often used in combination <a kid-*friendly* restaurant>

synonyms see **AMICABLE** — friend·li·ly \'fren(d)-lə-lē\ *adverb* — friend·li·ness *noun*

A more automated way to associate senses to lemma relies on the use of a lexical database like WordNet (WN). In WordNet, the senses are called *synsets* and a gloss (also called a short phrase), optionally with usage examples, is given for each synset. Moreover, as in MW, lexical relations (antonym, synonym, hyponym, etc.) are provided. In Table 2, we give WN entries (without lexical relations) of adjectives with frequency 9 or 8 in the resources files of Table 1.

If we consider the lemma 'friendly', MW and WN descriptions are not exactly the same but they share the personality-trait related senses: 1 = /pally/, 1a = /favorable/. This provides evidence that one can rely on MW and WN for a good coverage of the required senses. More important, the overall observation of the glosses related to personality description in MW and WN reveals that descriptions are generally given in terms of:

a) The *manner* a subject performs actions or activities: e.g. /alert/ = "quick and energetic"

b) The *attitude* of a subject interacting with others: e.g. /favorable/ = "inclined to help or support"

This is the reason why we propose in the next section a symbolic representation of the senses related to personality description in terms of Behavioral Schemes that capture subjects' manners and attitudes.

Table 2. WordNet synsets associated with the 5 most frequent p_i in the Cp corpus.

Freq	**Adjective**	**Synsets**	**Gloss** (non personality related synsets are in *small italic*)
9	friendly	/pally/	characteristic of or befitting a friend
		/allied/	*belonging to your own country's forces or those of an ally*
		/easy/	*easy to understand or use*
		/favorable/	inclined to help or support; not antagonistic or hostile
8	lively	/vital/	full of spirit
		/eventful/	*filled with events or activity*
		/frothy/	full of life and energy
		/Springy/	*elastic; rebounds readily*
		/alert/	quick and energetic
		/racy/	full of zest or vigor
8	kind	/tolerant/	tolerant and forgiving under provocation
		/genial/	agreeable, conducive to comfort
		/openhearted/	having or showing a tender and considerate and helpful nature; used especially of persons and their behavior
8	helpful	/helpful/	providing assistance or *serving a useful function*
8	ambitious	/pushy/	having a strong desire for success or achievement
		/challenging/	*requiring full use of your abilities or resources*

3. COMPUTATIONAL IMPLEMENTATION

From the examination of the relevant WN synsets (*i.e.* only those related to personality description) of the 25 most frequent adjectives, it is possible to establish that their glosses can generally be expressed as a Behavioral Scheme of the form: $\sigma \equiv F(P_i(a), ...)$ or $\sigma \equiv F(P_i(P_j(a), ...))$ where:

— F (performative): denotes a pair of notions (a positive one and its antonym) about a disjunction of P_i.

Ex: TEND (*resp.* AVOID): the subject tends to do or like (resp. avoid) $P_i(a)$;
POSSESS (*resp.* LACKOF): the subject has (resp. lacks) the feature(s) $P_i(a)$.

— P (predicate): denotes notions related to manner and attitudes of subject *x* about the entity a (optionally involving another subject *y*).

Ex: INTENT *x* intends to perform a in the near future;
ADOPT *x* adopts a as a goal for the near future;
SUGGEST *x* suggests *y* that action a should be a good thing to do;
TELL *x* tells *y* fact a.

Moreover, predicates can wrap a second order predicate P_j such as to express a modality.

Ex: ISQUICK(P_j(a)) *x* is quick to when doing $P_j(a)$.

— a (argument): denotes any entity that can be the object of a predicate.

Ex: operation any action in the system;
help-action *x* acts on the system on behalf of *y*;
help-information *x* provides *y* with a help information;
comfort *x* expresses "words of comfort" to *y*;
void/any are the empty argument and anything argument.

It is also useful to predefine *basic schemes* occurring frequently in the glosses. Basic schemes are noted as uppercase symbols defined with the following syntax: BASICSCHEME = F(P_i("a")).
For example, here is the description of some basic schemes (some are used in section 3.2):

```
HELPFUL         = TEND[INTENT["help-action"], SUGGEST["action"], TELL["help-information"]]
COMFORTING      = TEND[EXPRESS["comfort"]]
PAIR            = TEND[FEEL["same-rank"]]
JOYFUL          = TEND[FEEL["joy"]]
TOLERANT        = AVOID[REACTTO["provocation"]]
ENERGETIC       = POSSESS[PHYSICAL["energy"]]
SILLY           = LACKOF[ISSERIOUS["void"]]          // empty argument
```

Personality adjectives can be manually annotated by associating to each of their relevant WN synsets a set of elements, composed of schemes of the form F(P(...)) or of basic schemes symbols. For example, the annotations of adjectives 'friendly', 'lazy' and 'lively' are as follows:

```
friendly = {
    /pally/         = {HELPFUL, COMFORTING, TOLERANT, PAIR},    // use of predefined basic schemes
    /favorable/     = {HELPFUL}
}
lazy = {
    /workshy/       = AVOID[EXECUTE["action"], ISHARD["action"]]    // first scheme form
    /slow/          = TEND[ISSLOW[EXECUTE["action"]]]               // second scheme form (predicate wrapping)
}
lively = {
    /frothy/        = {"ENERGETIC"},
    /alert/         = TEND[ISQUICK[REACTTO["event"]]],
    /racy/          = TEND[ISQUICK[EXECUTE["action"]]]              // antonym of /slow/ = ISSLOW[EXECUTE["action"]]
}
```

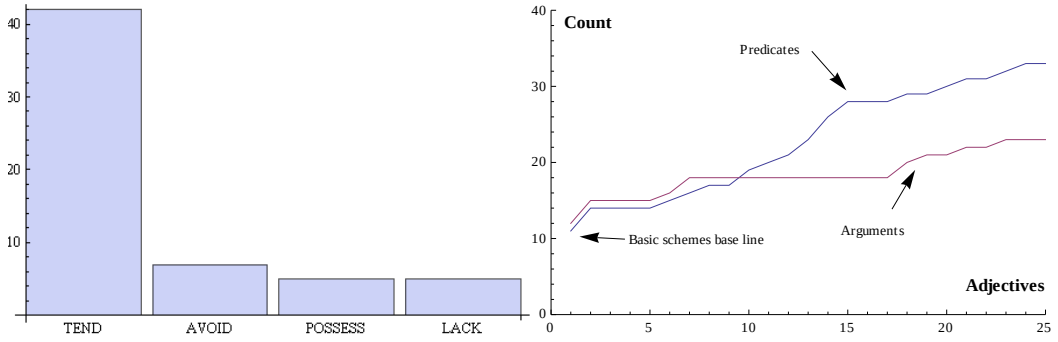

Figure 1. Frequency of the 4 main performatives (a) and evolution of the number of predicates and arguments used while annotating the adjectives (b), for the behavioral schemes of the 25 most frequent adjectives.

Table 3. Frequency of the 33 predicates and 23 arguments for the behavioral schemes of the 25 most frequent adjectives

Predicates	**Arguments**	**Freq.**
–	void	20
–	action	9
INTENT, FEEL	help-action	8
SUGGEST, TELL, EXPRESS	help-information	7
–	–	6
–	–	5
REACTTO, ISPLEASANT	comfort, provocation, joy, event	4
DESIRE, PHYSICAL, SHARE	same-rank, energy	3
OVERESTIMATE, OBEY, ISQUICK, ISEFFICIENT, EXECUTE, ISELEGANT, ISTENSE	success, positive-probability, command, goal, resources, joke	2
ADOPT, BUILD, ISDIGNIFIED, STRONG, ISWORKINGHARDON, ISAGITATED, TEND, ISSERIOUS, ISFOOLISH, ISSHOWY, ISOBTRUSIVE, ISACTIVE, ISSLOW, ISHARD, ISOK, EXPECT	plan, self-worth, emotion, thrilled, information, action-result, anxiety	1

From the annotation of the 25 first adjectives, a list of 57 synsets has been exhibited and defined thus producing 39 distinct behavioral schemes. Consequently, the two following observations can be made:

1) Figure 1a shows that the positive pole of the TEND operator is widely used to express personality traits; the POSSESS-LACK operator associated with features is much less used. Consequently, the proposed F operators have proved sufficient to express the 57 synsets associated with the 25 annotated adjectives.
2) Figure 1b shows that arguments are quite stable whereas the predicates are not yet in a log distribution with 25 adjectives: hence we can consider that we have a good coverage for the arguments but the coverage of the predicates is not yet complete.

In Table 3, we present the frequency of predicates and arguments used in the behavioral schemes found in the 25 most frequent adjectives. This distribution reveals that the most frequent predicates are either intrinsic (INTENT, FEEL, DESIRE) or interpersonal (SUGGEST, TELL, EXPRESS), and that frequent arguments (excluding 'void') are actions or are about actions, then about interaction (comfort, provocation).

CONCLUSION

In this study, we have collected a list of personality-trait adjectives from several thesaurus sources and extracted the lexical semantics associated with the most frequently used. Then a formal representation has been proposed to support the formal implementation of the semantics found in their related WordNet synsets.

Next work will be to annotate more adjectives to elicit more predicates so as to obtain a log curve in Figure 1.b. In the long term, main work will consist in the definition of a set of heuristics $H\sigma_i$ for the exhibited behavioral schemes. Meanwhile, using subsets of the behavioral schemes, it should be possible to put their $H\sigma_i$ to the test on a conversational agent system like the DIVA toolkit (Xuetao 2009). It will be used for experiments with human subjects in order to measure the extent and the precision of their actual perception of the behavioral schemes when they influence the rational process of the assistant agent.

4. REFERENCES

Allport, G.W. & Odbert, H.S., 1936. Trait names: a psycho-lexical study. Psychological Monographs, 47(211).

Anderson, N.H., 1968. Likableness ratings of 555 personality-trait words. Journal of Personality and Social Psychology, 9(3), 272-279

Baumgarten, F., 1933. Die Charaktereigenschaften [The character traits] A. Francke, ed. Beiträge zue Charakter-und Persönlichkeitsforschung, (1).

Bouchet F., Sansonnet J. P., Caractérisation de Requêtes d'Assistance à partir de corpus, MFI'07, Paris 2007

Cassell, J. et al. eds., 2000. Embodied Conversational Agents, MIT Press.

Ekman, P., Friessen, W.V. & Ellsworth, P.C., 1972. Emotion in the human face, Pergamon Press.

Fellbaum, C., 1998. WordNet: An Electronic Lexical Database, Cambridge, MA: MIT Press.

Goldberg, L.R., 1981. Language and individual differences: The search for universal in personality lexicons L. Wheeler, ed. Review of personality and social psychology, 2(1), 141-165.

Maes, P., 1994. Agents that reduce work and information overload. Commun. ACM, 37(7), 30-40.

McCrae, R.R. & Costa, P.T., 1987. Validation of the five-factor model of personality across instruments and observers. Journal of Personality and Social Psychology, 52(1), 81-90.

Russell, S. & Norvig, P., 2009. Artificial Intelligence: A Modern Approach 3 ed., Prentice Hall.

Sabouret N., Sansonnet J. P., Learning Collective Behavior from local Interactions, CEEMAS, Cracovia, 2001

Strapparava, C. & Mihalcea, R., 2008. Learning to identify emotions in text. In *Proceedings of the 2008 ACM symposium on Applied computing*. Fortaleza, Ceara, Brazil: ACM, pp. 1556-1560.

Strapparava, C. & Valitutti, A., 2004. WordNet-Affect: an Affective Extension of WordNet. In *Proc. of the 4th International Conference on Language Resources and Evaluation*. LREC. Lisbon, Portugal, pp. 1083-1086.

Valitutti, A., Strapparava, C. & Stock, O., 2004. Developing Affective Lexical Resources. *PsychNology Journal*, 2(1), 61-83.

Xuetao M., Bouchet F., J-P Sansonnet J.P., 2009. Intelligent Assisting Conversational Agents viewed through novice users' requests, International Conference Interfaces and Human Computer Interaction, IHCI 2009, Algarve, Portugal.