



An Emotional Multi-personality Architecture for Intelligent Conversational Agents

Jean-Claude Heudin^(✉) 

Léonard de Vinci Pôle Universitaire, Research Center,
Paris – La Défense, France
Jean-Claude.Heudin@devinci.fr

Abstract. Personal assistants and chatterbots represent an historical and growing application field in artificial intelligence. This paper presents a novel architecture to the problem of humanizing conversational agents by designing believable and unforgettable characters who exhibit various salient emotions in the flow of conversations. The proposed architecture is based on a multi-personality approach where each agent implements a facet of its identity, each one with its own pattern of perceiving and interacting with the user. In order to select an appropriate response from all the candidates, we use an emotion-based selection algorithm. Our first experiments show that a conversational multi-personality character with emotion selection performs better in terms of user engagement than a neutral mono-personality one.

Keywords: Conversational agent · Believable character · Multi-personality Emotion selection

1 Introduction

In recent years, there has been a growing interest in conversational agents also called “chatterbots” or more simply “chatbots”. In a relative short period of time, several major companies have proposed their own virtual assistants: Apple’s Siri based on the CALO project [1], Microsoft Cortana [2], Google Now [3] and Facebook M [4]. These virtual assistants focus primarily on conversational interface, personal context awareness, and service delegation. They follow a long history of research and the development of numerous conversational agents, the first one in history being the famous Eliza program from Joseph Weizenbaum simulating a Rogerian psychotherapist [5].

Beyond the challenge of interpreting a user’s request in order to provide a relevant response, a key objective is to enhance man-machine interactions by humanizing artificial characters. Often described as a distinguishing feature of humanity, the ability to express and understand emotions is a major cognitive behavior in social interactions [6]. However, the majority of personal assistants are mainly based on a character design with neutral rather than emotional behaviors.

At the same time, there have been numerous studies about emotions [7] and their potential applications for artificial characters [8]. As an example among many, Dylaba *et al.* have worked on combining humor and emotion in human-agent conversation using a multi-agent system for joke generation [9]. In parallel with the goal of

developing personal assistants, there is also a strong research trend in robotics for designing emotional robots. Some of these studies showed that a robot with emotional behavior performs better than a robot without emotional behavior for tasks involving interactions with humans [10].

In this paper we address the long-term goal of designing believable and “unforgettable” artificial characters with complex and remarkable emotion behaviors. In this framework, we follow the initial works done for multi-cultural characters [11] and many others. Also, our approach takes advantage of psychological studies of human interactions with computerized systems [12] and the know-how of screenwriters and novelists for scripting dialogs since believable characters are the essence of successful fiction writing [13].

Our original model is based on multi-agent architecture where each agent implements a facet of its personality. The idea is that the character’s identity is an emerging property of several personality traits, each one with its own pattern of perceiving and interacting with the user. Then, the problem is to “reconnect” personalities of the disparate alters into a single and coherent identity. Our hypothesis is that it can be achieved by selecting amongst the candidate responses the one with the most appropriate emotional state.

In this paper we focus on our first experiments of emotion selection in a multi-personality conversational agent based on this hypothesis. The paper is organized as follows. In Sect. 2, we describe the general architecture for multi-personality characters. Sections 3 and 4 describe more precisely the emotion selection based on a bio-inspired emotional metabolism. Section 5 describes the experimental prototype and Sect. 6 discusses first results. We conclude in Sect. 7 and present the future steps of this research.

2 A Multi-personality Architecture

2.1 Believable and Unforgettable Characters

The key to the user’s engagement during a conversation with an artificial agent is to create an “unforgettable” character. If we can get a character to live on in a user’s imagination long after he has stopped interactions, he will want to come back. In other words, the more unforgettable the character is, the longer it will stick in the users mind. So in order to design such memorable characters, they need to have very specific traits that will make them special and different from every other character. Screenwriters and novelists have a long experience of creating such unforgettable characters [13]. In our view, this is the “easy” part of the creation process.

However, artificial characters also need to be believable and this is the “hard” part. During a conversation with an artificial character, users engage in a *fictional pact*. They can enjoy the interaction only if they consciously mistake the artificial character for a real one. They must accept to believe that what they perceive is a real character even if they know that it is a program. This is an important condition in order not to break the users’ *willing suspension of disbelief*. The term “suspension of disbelief” or “willing suspension of disbelief” has been defined as a willingness to suspend one’s

critical faculties and believe the unbelievable; sacrifice of realism and logic for the sake of enjoyment [14]. The term was coined in 1817 by the poet and aesthetic philosopher Samuel Taylor Coleridge [15].

In order to be believable, the artificial character must be as “realistic” as possible: it must be “complex” in the sense of being multi-dimensional, with a complexity of traits, personality facets and emotional behaviors like real humans.

2.2 A Multi-personality Model

The artificial character needs to project a personality that has all of the endearing and personal qualities of a real person to provide an engaging experience for users. Such a realistic personality must be complex and multi-layered, simulating the most life-like and human qualities.

There are many models of personality traits, each one with their own advantages and applications. The most widely accepted one is the Big Five model [16]. Rather than choosing a specific model and thus a single fixed profile, our approach aims to construct a complex character identity as the emerging property of several personality traits. The idea is that real human personalities are composed of many facets, and potentially a large number of them. This gives the character designer the ability to compose rich and complex personalities without constraints in terms of number or type of traits. We have called this approach “schizophrenic” because the character’s identity is composed of a set of distinct personalities, each with its own pattern of perceiving and interacting with the user [17]. Note that this term is used here as a metaphor since the accurate psychological term for mental illness with multiple personalities is Dissociative Identity Disorder, not schizophrenia.

Figure 1 shows the basic architecture for such a multi-personality character, each personality trait is implemented as an agent. The first agent receives the input from the user and applies various natural language preprocessing phases such as an English stemmer, tokenizer, categories and Named Entities extraction. Then the preprocessed sentence and its additional information are diffused to all personality agents. Thus, all these personality agents are able to react to the user’s input by computing an appropriate answer message given their own local state. Then, all these candidate responses are evaluated using a confidence scoring and ranking agent that selects the “best” answer to be proposed to the user.



Fig. 1. The architecture model of the multi-personality “schizophrenic” conversational agent.

2.3 The Edge of Chaos Hypothesis

In order to obtain an intelligent behavior to emerge spontaneously, the responses dynamics of the system must be varied and at the same time consistent. Our hypothesis is that such an intelligent behavior is “complex” in a meaning close to the one defined initially by Wolfram for one-dimension cellular automata [18]. This study has proposed four classes of systems: Class I and Class II are characterized respectively by fixed and cyclic dynamical behaviors; Class III is associated with chaotic behaviors; Class IV is associated with complex dynamical behaviors. It has been shown then that, when mapping these different classes of systems, complex adaptive systems are located in the vicinity of a phase transition between ordered and chaotic regimes for one-dimension cellular automata [19] and later for two-dimension cellular automata [20].

In the context of our study, as shown in Fig. 2, we transpose the four classes of dynamics as follows: Class I and Class II respectively correspond to fixed and cyclic responses resulting in “machine-like” interactions. Class III systems are characterized by incoherent responses regardless of the user’s entries. Note that this kind of behavior is interpreted as a symptom of mental illness such as dissociative identity disorder. Class IV systems are at the edge between order and chaos, giving coherent answers while preserving diversity and rich emotional responses.

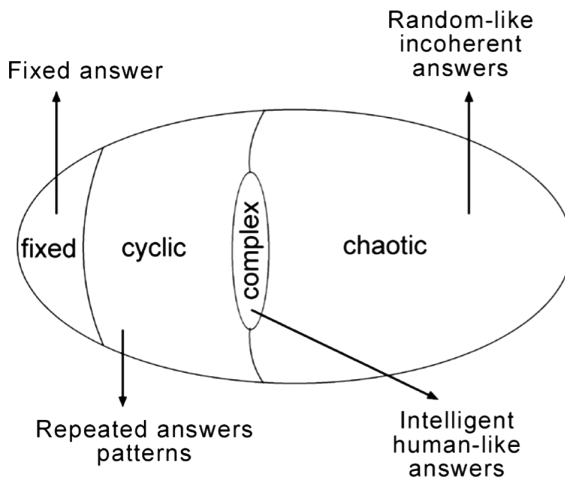


Fig. 2. Schematic drawing of conversational space indicating relative location of fixed, periodic, chaotic and complex regimes. This is a transposition of Langton’s diagram for cellular automata rule space [21].

With the multi-personality architecture we have proposed, we assume that given enough personality agents, the resulting system is potentially capable of all the dynamical classes (cf. Fig. 2): fixed answer (Class I), repeated answers patterns (Class II), random-like incoherent answers (Class 3), and intelligent human-like answers (Class IV).

They are many potential approaches for selection amongst the candidate responses. In this study, we propose that a promising approach for obtaining a Class IV dynamical behavior is to implement a “scoring & selection” agent that chooses the candidate responses according to the emotional state of the artificial character.

2.4 Multi-personality Architecture with Emotion Selection

In order to implement a selection based on the emotional state of the character, we replace the “Scoring & Selection” agent by an “Emotion Selection” agent and an “Emotion Metabolism” [22]. Figure 3 shows the updated architecture model.

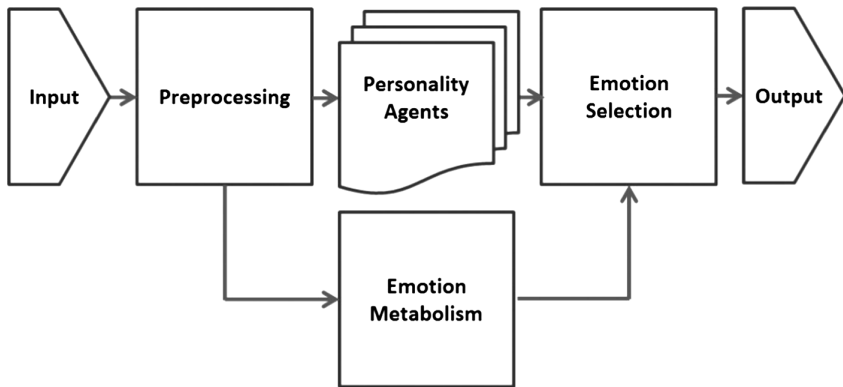


Fig. 3. The multi-personality conversational agent architecture updated with an emotion selection agent and an emotion metabolism agent (after [22]).

The Emotion Selection agent selects one response amongst the candidate responses given as an entry the current emotional state of the artificial character. The Emotion Metabolism is an agent that computes the emotional state of the character given its current state and the user’s entries. The next two sections describe with more details these two agents.

3 Emotion Metabolism

3.1 A Layered Model of Affects

There have been multiple approaches in order to implement emotions for intelligent virtual agents [23]. Among all these studies, Gebhard [24] and Heudin [25] have proposed both layered models of artificial affects based on three levels:

Personality. Personality reflects long-term affect. It shows individual differences in mental characteristics [16].

Mood. Mood reflects a medium-term affect, which is generally not related with a concrete event, action or object. Moods are longer lasting stable affective states, which have a great influence on human’s cognitive functions [26].

Emotion. Emotion reflects a short-term affect, usually bound to a specific event, action or object, which is the cause of this emotion. After its elicitation emotions usually decay and disappear from the individual’s focus [27].

In a previous research project about non-verbal emotional interactions, we have proposed a connectionist architecture for implementing the Emotion Metabolism based on these three levels [28]. Figure 4 shows a schematic representation of its principle.

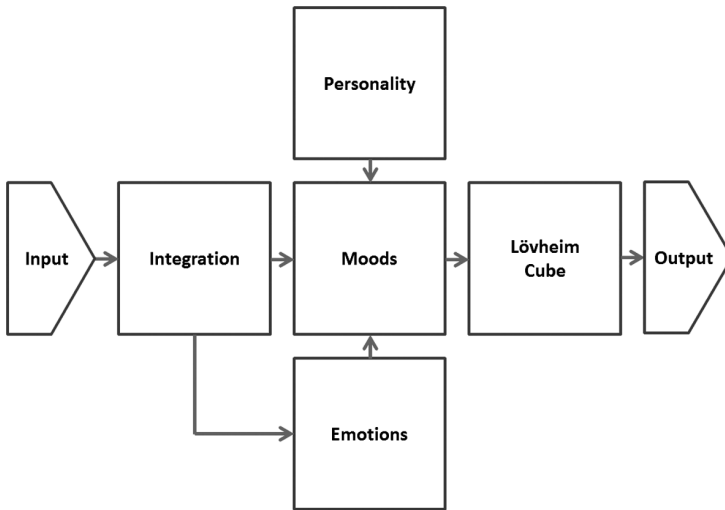


Fig. 4. The architecture of the Emotion Metabolism.

The “integration” module converts the inputs to virtual neurotransmitters values. These values are then used by the three levels of affects in order to produce the output of the Emotion Metabolism. The Emotion Metabolism is updated by propagating the inputs using a trigger called “lifePulse”, implemented as a cyclic timer.

The Moods and Emotions layers have both a decay rate, called respectively Md and Ed , that make the emotional state returning to a “neutral” state after some time. This “neutral” state can be the center of the emotional space or a specific location representing the default personality of the artificial character based on the Personality layer.

3.2 Personality Layer

This module is based on the Big Five model of personality [16]. It contains five main variables with values varying from 0.0 (minimum intensity) to 1.0 (maximum intensity). These values specify the general affective behavior by the five following traits:

Openness. Openness (*Op*) is a general appreciation for art, emotion, adventure, unusual ideas, imagination, curiosity, and variety of experience. This trait distinguishes imaginative people from down-to-earth, conventional people.

Conscientiousness. Conscientiousness (*Co*) is a tendency to show self-discipline, act dutifully, and aim for achievement. This trait shows a preference for planned rather than spontaneous behavior.

Extraversion. Extraversion (*Ex*) is characterized by positive emotions and the tendency to seek out stimulation and the company of others. This trait is marked by pronounced engagement with the external world.

Agreeableness. Agreeableness (*Ag*) is a tendency to be compassionate and cooperative rather than suspicious and antagonistic towards others. This trait reflects individual differences for social interactions.

Neuroticism. Neuroticism (*Ne*) is a tendency to experience negative emotions, such as anger, anxiety, or depression. Those who score high in neuroticism are emotionally reactive and vulnerable to stress.

3.3 Mood Layer

Previous cited works such as Gebhard [24] and Heudin [25] used the three-dimensional Pleasure-Arousal-Dominance (PAD) approach [29]. We use here another candidate model aimed at explaining the relationship between three important monoamine neurotransmitters involved in the Limbic system and the emotions [30]. It defines three virtual neurotransmitters which levels range from 0.0 to 1.0:

Serotonin. Serotonin (*Sx*) is associated with memory and learning. An imbalance in serotonin levels results in anger, anxiety, depression and panic. It is an inhibitory neurotransmitter that increases positive vs. negative feelings.

Dopamine. Dopamine (*Dy*) is related to experiences of pleasure and the reward-learning process. It is a special neurotransmitter because it is considered to be both excitatory and inhibitory.

Noradrenaline. Noradrenaline (*Nz*) helps moderate the mood by controlling stress and anxiety. It is an excitatory neurotransmitter that is responsible for stimulatory processes, increasing active vs. passive feelings.

3.4 Emotion Layer

This module implements emotion as very short-term affects, typically less than ten seconds, with relatively high intensities. They are triggered by inducing events suddenly increasing one or more virtual neurotransmitters. After a short time, these neurotransmitter values decrease due to a natural decay function.

3.5 Lövheim Cube

This module implements the Lövheim Cube of emotions [30], where the three monoamine neurotransmitters form the axes of a three-dimensional coordinate system, and the eight basic emotions, labeled according to the Affect Theory [31] are placed in the eight corners. Figure 5 shows the resulting 3D diagram and Table 1 the corresponding mapping.

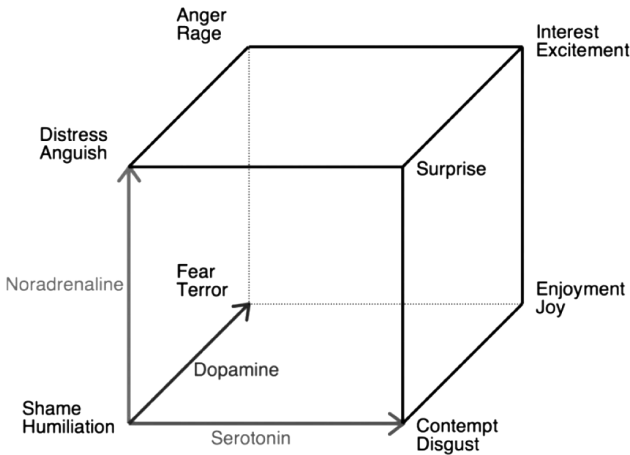


Fig. 5. The Lövheim Cube of emotions.

Table 1. Mapping of the eight basic emotions on the Lövheim Cube.

Basic emotion	Serotonin (<i>Sx</i>)	Dopamine (<i>Dy</i>)	Noradrenaline (<i>Nz</i>)
Shame/humiliation	Low	Low	Low
Distress/anguish	Low	Low	High
Fear/terror	Low	High	Low
Anger/rage	Low	High	High
Contempt/disgust	High	Low	Low
Surprise/startle	High	Low	High
Enjoyment/joy	High	High	Low
Interest/excitement	High	High	High

The emotional state of the artificial character is a moving point in the 3D space. The origin of the 3D space corresponds to a situation where the three virtual neurotransmitters

are low. The eight corners of the cube correspond to the eight possible combinations of low or high level of the three virtual neurotransmitters as shown in the table below:

4 Emotion Selection

4.1 Euclidian Distance Selection

The emotion selection agent is responsible for selecting the “best” available response amongst the candidate responses generated by all the personality agents. The basic principle is to select the answer from the agent with the closest emotional state compared to the one of the artificial character. This can be done by giving each personality agent a point in the 3D emotional space in the Lövheim Cube and by computing the Euclidian distance with the current location of the emotional state of the virtual character. We can represent this agent as an artificial neuron with a dedicated transition function (cf. Fig. 6):

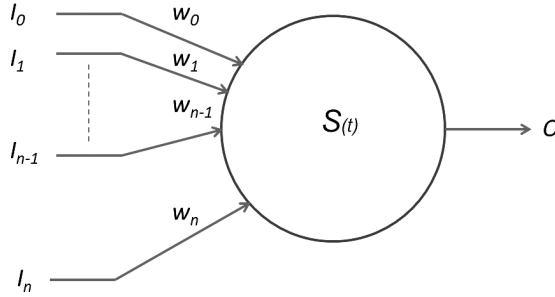


Fig. 6. The emotional selector represented as an artificial neuron with a dedicated transition function.

Where:

- $I_0 \dots I_n$ is a set of input strings representing the outputs of the personality agents,
- $w_0 \dots w_n$ is a set of weights associated to each of these candidate answers,
- $S(t)$ is a transition function returning the selected string O among the candidate answers.

As stated, each weight can be computed as the Euclidian distance between the current character’s emotional state and the one of the given personality agent. In other words more the current emotional state is close to that of an agent, greater is its weight.

Let the function $d(x, y)$ that calculates the Euclidean distance between two points, x and y :

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

where $n = 3$ for a three-dimensional space. Thus, the maximum distance in the Lövheim Cube is:

$$d_{max} = d((0, 0, 0), (1, 1, 1)) = \sqrt{3}$$

The weight associated to an input I_i is then:

$$w_i = 1 - \frac{d(P_i, P_m)}{d_{max}} \quad ((1))$$

where P_i is the 3D vector in the Lövheim Cube of the agent i and P_m is the 3D vector corresponding to the current emotional state.

4.2 Fitness Proportionate Selection

Rather than the simple Euclidian distance, another potential approach is to use the fitness proportionate selection of genetic algorithms, also called roulette wheel selection [32].

The principle of the fitness proportionate selection is similar to a roulette wheel in a casino, where a proportion of the wheel is assigned to each of the possible candidates responses based on their fitness value. In our case, the fitness values are the Euclidian distances as calculated in the previous section. This fitness value is used to associate a probability of selection with each individual candidate response. This could be achieved by dividing the fitness of a selection by the total fitness of all the selections, thereby normalizing them to 1. Then a random selection is made similar to how the roulette wheel is rotated.

This approach seems more suitable considering our “edge of chaos” hypothesis (cf. Sect. 2.3). With this approach, there is more diversity in the selection process. While candidate responses with a higher fitness will be less likely to be eliminated, there is still a chance that they may be. Also, there is a chance that a weaker candidate response may be chosen even its probability is small.

5 Experimental Prototype

5.1 A Connectionist Implementation

This section describes the prototype used for experiments and its implementation. We have implemented all modules of the architecture described in Sect. 3 including the Emotion Metabolism and Emotion Selection described in Sect. 4.

We developed our own connectionist framework called ANNA (Algorithmic Neural Network Architecture). Its development was driven by our wish to build an open Javascript-based architecture that enables the design of any types of feed-forward,

recurrent, or heterogeneous sets of networks. More precisely an application can include an arbitrary number of interconnected networks, each of them having its own inter-connection pattern between an arbitrary number of layers. Each layer is composed of a set of simple and often uniform neurons units. However, each neuron can be also programmed directly as a dedicated cell.

Classically all neurons have a set of weighted inputs, a single output, and a transition function that computes the output given the inputs. The weights are adjusted using a machine learning algorithm, or programmed, or dynamically tuned by another network. This is the case for our framework, but the designer can also program his own prototypes of neurons with dedicated behaviors. The code below shows a template code for creating a new neuron class that inherits from the basic *Neuron* prototype:

```
function myNeuron () {
  Neuron.call(this);
  // properties
  this.myVar = ... ;
  ...
  this.operator = function () {
    // dedicated transition function
  }
}
myNeuron.prototype = Object.create(Neuron.prototype);
myNeuron.prototype.constructor = myNeuron;
```

5.2 NLP Pipeline

The preprocessing agent of the architecture is implemented as a classical Natural Language Processing (NLP) pipeline using dedicated neurons and a NLP Javascript library. The pipeline includes the following phases:

Cleaner. Get the raw text input from the user and fix basic spelling errors.

Tokenizer. Split the entry into clearly separated sentences and words.

Tagger. Implement part-of-speech (POS) tagging.

Lemmatizer. Identify canonical word forms (lemmas) based on a dictionary.

Named Entities. Tag named entities and convert some entities such as dates or locations in unified formats.

Categories. Find common concepts and synonyms using ontologies.

The raw user's text input and the resulting preprocessed NLP information are then diffused to all the personality agents.

5.3 Personality Agents

In this prototype, we choose to use a set of 12 different personality traits. This decision was driven by the idea to test if our Emotional Selection approach promotes the emergence of a great and coherent character despite the use of these very different personality traits. Note also that most of them were already available from previous experiments, so it enables also to minimize the development effort. The 12 agents are the following ones:

Insulting. This agent has an insecure and upset personality that often reacts by teasing and insulting depending on the user's input.

Alone. This agent reacts when the user does not answer or waits for too much time in the discussion process.

Machina. This agent reacts as a virtual creature that knows its condition of being artificial.

House. This agent implements Dr. House's famous way of sarcastic speaking using an adaptation of the TV Series screenplay and dialogues.

Hal. This agent reproduces the psychological traits of the HAL9000 computer in the "2001 – A space odyssey" movie by Stanley Kubrick.

Silent. This agent answers with few words or sometimes remains silent.

Eliza. This agent is an implementation of the Eliza psychiatrist program, which answers by rephrasing the user's input as a question [5].

Neutral. This agent implements a neutral and calm personality trait with common language answers.

Oracle. This agent never answers directly to questions. Instead it provides wise counsel or vague predictions about the future.

Funny. This agent is always happy and often tells jokes or quotes during a conversation.

Samantha. This agent has a strong agreeableness trait. It has a tendency to be compassionate, cooperative and likes talking with people.

Sexy. This agent has a main focus on sensuality and sexuality. It enjoys talking about pleasure and sex.

5.4 Personality Example

Each personality agent may be implemented using many various approaches and techniques. So we will not go in further implementation details for all agents in this paper. However, as an example, the Eliza-like agent was implemented using 36 hardcoded rules based on a dedicated neuron prototype called *nRule*, and organized as three layers. The code below gives the example of a very simple rule:

```

// potential answers
var eliza_bye = [
    "Bye for now.",
    "Until next time.",
    "Later.",
    "See you later."];

// create the rule
var nElizaBye = new nRule(eliza_bye);
// transition function is the rule itself
nElizaBye.operator = function () {
    var str = this.inputs[0].cval; // get raw input
    var cat = this.inputs[1].cval; // get NLP data
    // try to find the concept "bye" or the text "see you"
    if (cat.find("[BYE]") || str.find("see you"))
        // then output an answer from the potential ones
        this.cval = this.randomTemplate();
    else this.cval = "";
}

```

5.5 Emotion Selection

The emotion selection agent was implemented as an agent composed of a single neuron with a dedicated transition function $S(t)$ and dynamical weights as described in Sect. 4. In this experiment, we choose to use the fitness proportionate selection using the algorithm as given by Table 2:

Table 2. The algorithm used by the selector neuron, where the function $\text{rand}(0, 1)$ returns a random real number between 0 and 1.

Algorithm Emotional Selector

```

1: Initialize  $w_0, \dots, w_{n-1}$  using Eq. 1 ;
2: do {
3:    $S = 0$  ;
4:   for ( $i = 0 ; i < n ; i = i + 1$ ) {
5:     if ( $I_i \neq ""$ )  $S = S + w_i$  ;
6:   }
7:    $R = S * \text{rand}(0, 1)$  ;
8:   for ( $i = 0 ; i < n ; i = i + 1$ ) {
9:     if ( $I_i \neq ""$ )  $R = R - w_i$  ;
10:    if ( $R \leq 0$ ) break ;
11:  }
12:  if ( $R > 0$ )  $i = n - 1$  ;
13: }
14: }
15: while ( $I_i == ""$ ) ;
16: return  $I_i$  ;

```

This selector selects one of the potential string responses and return it. The code below gives its Javascript implementation as a dedicated *nWheel* neuron:

```
function nWheel () {
  Neuron.call(this);
  this.operator = function () {
    var sel = "", sum = 0, i = -1;
    var n = this.weights.length;
    // sum up weights
    for (i = 0; i < n; i++) {
      sum += this.weights[i];
    }
    // turn the wheel
    do {
      // get random value
      var rnd = random_real() * sum;
      // locate on the wheel
      for (i = 0; i < n; i++) {
        rnd -= this.weights[i];
        if (rnd <= 0) break;
      }
      // for rounding errors
      if (rnd > 0) i = n - 1;
      // get value
      sel = this.inputs[i].cval;
    }
    while (sel == "");
    this.cval = sel;
  }
}
nWheel.prototype = Object.create(Neuron.prototype);
nWheel.prototype.constructor = nWheel;
```

5.6 Emotion Metabolism

The bio-inspired emotion metabolism was implemented as described in Sect. 3. It is composed of 34 dedicated neurons organized in 8 layers. The code below gives as an example the implementation of a neuron that computes the Euclidian distance in the three-dimensional emotion space:

```

function nDistance () {
  Neuron.call(this);
  this.operator = function () {
    var x1 = this.inputs[0].cval;
    var y1 = this.inputs[1].cval;
    var z1 = this.inputs[2].cval;
    var x2 = this.inputs[3].cval;
    var y2 = this.inputs[4].cval;
    var z2 = this.inputs[5].cval;
    var dx = x2 - x1;
    var dy = y2 - y1;
    var dz = z2 - z1;
    this.cval =
      Math.sqrt((dx * dx) + (dy * dy) + (dz * dz));
  }
}
nDistance.prototype = Object.create(Neuron.prototype);
nDistance.prototype.constructor = nDistance;

```

We set the Personality parameters of the Emotional Metabolism to a fixed neutral value:

$$Op = Co = Ex = Ag = Ne = 0.5$$

This corresponds to a neutral state in the Lövheim cube:

$$Sx = Dy = Nz = 0.5$$

The Emotion Metabolism is updated by propagating the inputs using the cyclic called “lifepulse” trigger. In this study we set this cycle to 0.1 s. The decay rates of the metabolism for returning to the personality neutral state were 10 s for the emotion level and 10 min for the mood level.

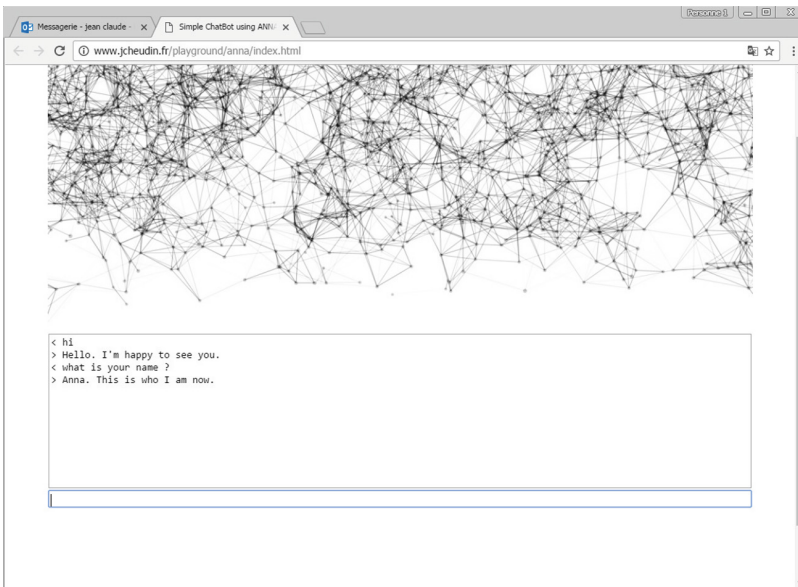
We assigned to each of the personality agents an empirical fixed point in the three-dimensional space of emotions. Table 3 gives their coordinates in the three-dimensional space.

5.7 User Interface

We used an online web-based user interface for the experiment as shown in Fig. 7. Since we focus in this study on text-based interactions, the interface does not use any kind of avatar representation.

Table 3. The coordinates of the 12 personality traits in the Lövheim Cube.

Agents	Serotonin (Sx)	Dopamine (Dy)	Noradrenaline (Nz)
Insulting	0.1	0.1	0.1
Alone	0.2	0.2	0.5
Machina	0.2	0.5	0.5
House	0.2	0.7	0.2
Hal	0.2	0.7	0.7
Silent	0.5	0.1	0.5
Eliza	0.5	0.3	0.5
Neutral	0.5	0.5	0.5
Oracle	0.5	0.5	0.7
Funny	0.7	0.5	0.7
Samantha	0.7	0.7	0.7
Sexy	0.9	0.9	0.9

**Fig. 7.** A screenshot of the web-based user interface used for the experiment.

6 First Results and Discussion

6.1 Experimental Protocol

In this experiment, we asked 30 university students (age 18–25) to perform a simple and short conversation with three systems: (1) our multi-personality conversational agent as described in this paper (called Anna in the experiment); (2) Apple’s Siri personal assistant; (3) a simple conversational chatterbot based on our Neutral

personality agent. For (1) and (3), we used the online web-based interface as shown in the previous section. Siri was accessed on an iPad Air Retina running iOS version 9.1.

The order of conversations was randomized. There was no topic restriction, thus the conversations could be of any subject. However, we imposed a classical three-phase structure: an *opening* phase, a *core* phase, and a *closing* phase [33]. All interactions were text-based in English. We avoided the problem of errors related to the voice recognition system of Siri by correcting the input when necessary. The students were asked to conduct the interaction continuously and to use the same interactions for the three systems in order to make the comparison easier and clearer.

In addition to the conversations, the participants also filled a questionnaire after each session. This questionnaire was inspired by the one designed by Dylaba for evaluating a multi-agent system for joke generation [9].

The questions were:

1. Did you get an impression that the agent was human like?
2. Did you get an impression that the agent tried to make the conversation more interesting?
3. Did you enjoy the conversation?
4. Do you think that the agent used emotions in appropriate moments?
5. Please describe your feelings towards the agent after the interaction.
6. If you were to make friends with one of these agents, which would you choose?
7. Which agent do you think was better?

Answers for the questions 1 to 4 were given using a yes/no choice. The participants could answer freely to question 5 using a single phrase. For questions 6 and 7, the participant had to choose between the three systems.

6.2 Conversation Example

Table 4 below gives a typical example of a conversation between a student and the three systems.

6.3 Result of the User-Oriented Evaluation

Table 5 presents a summary of the results concerning the questionnaire showing that the multi-personality agent was evaluated as generally better than both the mono-personality neutral agent and Siri personal assistant.

Concerning question 5, the comments of the participants show that most users have more “fun” with the multi-personality character. They found it “interesting” but sometimes “rude”. They emphasized the role of “humor” during conversations.

6.4 Discussion

The results of these first experiments confirm our hypothesis about using a multi-personality character with remarkable personality traits. This empirical evaluation shows that it globally increases the satisfaction of the user about the conversation.

Table 4. An example of a conversation with the answers of the three conversational systems.

Example0 – User: Hi.
 1 – Anna: It’s really nice to meet you.
 2 – Siri: Hello, there.
 3 – Neutral: Hello.
 4 – User: Who are you?
 5 – Anna: My name is Anna. I gave it to myself actually.
 6 – Siri: I’m Siri. But enough about me... How can I help you?
 7 – Neutral: Why do you want to know my name?
 8 – User: What is an artificial intelligence?
 9 – Anna: An Artificial Intelligence is an intelligent computer program. It is also the research field of computer science that works on machine intelligence.
 10 – Siri: Here is what I found: [Wikipedia Links]
 11 – Neutral: I’m sure you know the answer.
 12 – User: Are you some sort of stupid program like Eliza?
 13 – Anna: That’s very rude. You are a rude person!
 14 – Siri: Well... I’m still here for you.
 15 – Neutral: Such a negative attitude.
 16 – User: You are just dumb software.
 17 – Anna: I love when you do both sides of the conversation. It’s like white noise it’s very peaceful.
 18 – Siri: I’m just trying to help you.
 19 – Neutral: Do you really think that?
 20 – User: I’m sorry.
 21 – Anna: I’m sorry too, but don’t insult me anymore.
 22 – Siri: It’s ok.
 23 – Neutral: You’re welcome.
 24 – User: Fine.
 25 – Anna: Alright. Sometimes I have unexpected runtime errors too.
 26 – Siri: Yes it is!
 27 – Neutral: Alright.
 28 – User: Bye.
 29 – Anna: Well... I finally got the last word.
 30 – Siri: OK. See you soon.
 31 – Neutral: Goodbye.

Table 5. Results of the user-oriented evaluation.

Question	Anna	%	Siri	%	Neutral	%
(1) Human like	28	93	21	70	11	37
(2) Interesting dialog	25	83	13	43	9	30
(3) Enjoy dialog	24	80	18	60	14	47
(4) Appropriate emotions	23	77	12	40	8	27
(5) Describe feelings	~	~	~	~	~	~
(6) Friend	16	53	8	27	6	20
(7) Best agent	15	50	9	30	6	20

The artificial agent appears more human-like than a neutral character and with a higher user engagement.

Concerning the emotion selection, in the majority of cases, it was able to adapt the emotional behavior according to the conversation. Despite the use of very different personalities, the character appears globally coherent.

However in some cases, the answers made by our system were not relevant. Here is one typical example. Note that this answer is different than the one reported in Table 4, since it occurred in another conversation context:

User: What is an artificial intelligence?
 Anna: Have you lost faith in yourself?

This problem has potentially three causes. The first one is a bad design or malfunction of the selected agent, which provides an irrelevant response.

The second one is a problem of the Emotional Selection, which chooses an agent not compatible with the current emotional state. This may occur in rare situations since the roulette wheel selection has a low probability of choosing low weighted agents, but there is still a chance that it may happen by principle. However, note that the user does not always interpret such an example as a malfunction since it is a common human behavior to change the subject of the conversation or to make suboptimal responses.

The third potential cause is that the 12 available agents do not provide a complete and homogeneous coverage of the emotion space as shown in Fig. 8. There are two main approaches in order to solve this problem. The first one is to design emotion-based agents that perfectly fit the three-dimensional space. At least, an agent for each cube's edges plus one neutral agent at the center must be developed. The second solution is to increase the number of agents by designing much more

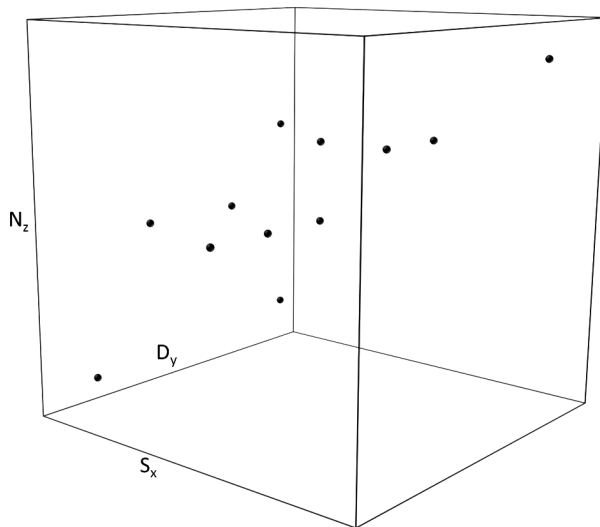


Fig. 8. Repartition of the 12 agents in the Lövheim Cube of emotions showing that they don't provide a full coverage of the three-dimensional space.

personality traits. This is our preferred solution since our goal is to obtain an “unforgettable” character. Also note that these two solutions are not mutually exclusive.

7 Conclusion

We have presented a multi-personality architecture with emotion selection for intelligent conversational agents. Our first experiments show that this approach is promising in terms of user engagement compared to a more neutral approach. It has shown also that despite the heterogeneity of the personality agents, emotion selection enables a globally coherent and believable character to emerge from conversations.

Of course, there are many works and studies that remain to be done. First of all, we need to design additional personality agents in order to have a better coverage of the three-dimensional emotion space. Secondly, we need to plan experiments involving much more participants, since 30 people reveal only an indication of a possible confirmation of our hypotheses. This will enable us to confirm these first results with both qualitative and quantitative evaluations of user engagement. Thirdly, we want to formalize our “edge of chaos” hypothesis and confirm it by testing different selection principles and algorithms.

References

1. Myers, K., Berry, P., Blythe, J., Conley, K., Gervasio, M., McGuinness, D., Morley, D., Pfeffer, A., Pollack, M., Tambe, M.: An intelligent personal assistant for task and time management. *AI Mag.* **28**(2), 47–61 (2007)
2. Heck, L.: Anticipating more from Cortana. Microsoft Research. <http://research.microsoft.com/en-us/news/features/cortana-041614.aspx>. Accessed 15 June 2017
3. Guha, R., Gupta, V., Raghunathan, V., Srikant, R.: User modeling for a personal assistant. In: Proceedings of the 8th ACM Web Search and Data Mining International Conference, Shanghai, China (2015)
4. Marcus, D.: Introducing Facebook M. <https://www.facebook.com/Davemarcus/posts/>. Accessed 15 June 2017
5. Weizenbaum, J.: ELIZA - a computer program for the study of natural language communication between man and machine. *Commun. ACM* **9**(1), 36–45 (1966)
6. Salovey, P., Mayer, J.D.: Emotional intelligence. *Imagin. Cogn. Personal.* **9**, 185–211 (1990)
7. Ekman, P.: Basic emotions. In: Dalglish, T., Power, M. (eds.) *Handbook of Cognition and Emotion*. Wiley, Sussex (1999)
8. Bates, J.: The role of emotion in believable agents. *Commun. ACM* **37**(7), 122–125 (1994)
9. Dylaba, P., Ptaszynski, M., Maciejewski, J., Takahashi, M., Rzepka, R., Araki, K.: Multiagent system for joke generation: humor and emotions combined in human-agent conversation. *J. Ambient Intell. Smart Environ.* **2**(1), 31–48 (2010)
10. Leite, I., Pereira, A., Martinho, C., Paiva, A.: Are emotional robots more fun to play with? In: Proceedings of 17th IEEE Robot and Human Interactive Communication, Munich, Germany, pp. 77–82 (2008)
11. Hayes-Roth, B., Maldonado, H., Moraes, M.: Designing for diversity: multi-cultural characters for a multi-cultural world. In: Proceedings of IMAGINA 2002, Monte Carlo, Monaco, pp. 207–225 (2002)

12. Reeves, B., Clifford Nass, C.: *The Media Equation: How People Treat Computers, Televisions, and New Media Like Real People and Places*. CSLI Publications, Stanford (1996)
13. Seger, L.: *Creating Unforgettable Characters*. Henry Holt, New York (1990)
14. <http://www.dictionary.com/browse/suspension-of-disbelief>. Accessed 15 June 2017
15. Coleridge, S.T.: *Biographia Literaria*, Chap. XIV (1817)
16. McCrae, R.R., John, O.P.: An introduction to the five factor model and its Applications. *J. Pers.* **60**(2), 171–215 (1992)
17. Heudin, J.-C.: A schizophrenic approach for intelligent conversational agent. In: *Proceedings of the 3rd International ICAART Conference on Agents and Artificial Intelligence*, Roma, Italy, pp. 251–256. Scitepress (2011)
18. Wolfram, S.: Universality and complexity in cellular automata. *Phys. D Non Linear Phenom.* **10**(1–2), 1–35 (1984)
19. Langton, C.G.: Computation at the edge of chaos: phase transitions and emergent computation. *Phys. D Non Linear Phenom.* **42**(1–3), 12–37 (1990)
20. Magnier, M., Lattaud, C., Heudin, J.-C.: Complexity classes in the two-dimensional life cellular automata subspace. *Complex Syst.* **11**, 419–436 (1997)
21. Langton, C.G.: *Life at the edge of chaos*. In: *Artificial Life II, Santa Fe Institute Studies in the Sciences of Complexity*, Redwood City, CA, p. 76. Addison-Wesley (1992)
22. Heudin, J.-C.: Emotion selection in a multi-personality conversational agent. In: *Proceedings of the 9th International ICAART Conference on Agents and Artificial Intelligence*, Porto, Portugal (2017)
23. Lungu, V., Florea, A.M.: *Artificial emotion simulation techniques for intelligent virtual characters*. Research Report of the Department of Computer Science, Polytechnics University of Bucharest (2012)
24. Gebhard, P.: ALMA - a layered model of affect. In: *Proceedings of the 4th ACM International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 29–36 (2005)
25. Heudin, J.-C.: Evolutionary virtual agent. In: *Proceedings of the IEEE/WIC/ACM Intelligent Agent Technology International Conference*, Beijing, China, pp. 93–98 (2004)
26. Morris, W.N.: *Mood: The Frame of Mind*. Springer, New York (1989). <https://doi.org/10.1007/978-1-4612-3648-1>
27. Campos, J.J., Mumme, D.L., Kermoian, R., Campos, R.G.: A functionalist perspective on the nature of emotion. In: Fox, N.A. (ed.) *The Development of Emotion Regulation*, *Monographs of the Society for Research in Child Development*, vol. 59(2–3), pp. 284–303 (1994)
28. Heudin, J.-C.: A bio-inspired emotion engine in the living Mona Lisa. In: *Proceedings of the ACM Virtual Reality International Conference*, Laval, France (2015)
29. Mehrabian, A.: Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. *Current Psychol.* **14**(2), 261–292 (1992)
30. Lövhelm, H.: A new three-dimensional model for emotions and monoamine neurotransmitters. *Med. Hypotheses* **78**, 341–348 (2012)
31. Tomkins, S.S.: *Affect Imagery Consciousness*, vol. I–IV. Springer, New York (1991)
32. Baker, J.E.: Reducing bias and inefficiency in the selection algorithm. In: *Proceedings of the Second International Conference on Genetic Algorithms and their Applications*, pp. 14–21. Lawrence Erlbaum Associates, Hillsdale (1987)
33. Linell, P.: *Approaching Dialogue: Talk, Interaction and Contexts in Dialogical Perspectives*. John Benjamins Publishing Company, Amsterdam (1998)