Cite this paper:
Galindo, M.J.O., Moreno, L.A.M., Rojas-Velázquez, D., Nieves, J.C. (2021). E-Friend: A Logical-Based AI Agent System Chat-Bot for Emotional Well-Being and Mental Health. In: Sarkadi, S., Wright, B., Masters, P., McBurney, P. (eds) Deceptive AI. DeceptECAI DeceptAI 2020 2021. Communications in Computer and Information Science, vol 1296. Springer, Cham. https://doi.org/10.1007/978-3-030-91779-1_7

E-friend: A logical-based AI Agent System Chat-bot for Emotional Well-being and Mental Health

Mauricio J. Osorio Galindo[1], Luis A. Montiel Moreno[2 [0000−0003−0276−0559]], David Rojas-Velázquez,[3 [0000−0003−4402−4736]], and Juan Carlos Nieves,[4 [0000−0003−4072−8795]]

[1] Universidad de las Américas Puebla, Ex Hacienda Sta. Catarina Mártir S/N. San Andrés Cholula, Puebla 72810, MX - mauricioj.osorio@udlap.mx
[2] Independent Researcher - luis.montiel.moreno@gmail.com
[3] Independent Researcher - davidrojasvelazquez@gmail.com
[4] Department of Computing Science, Umeå University, Sweden - jcnieves@cs.umu.se

Abstract. In this work, it is proposed the design of a Reasoning Logical Based Intelligent Agent System Chat-bot for Dialogue Composition (DC) named E-friend, which uses Logic Programming (LP) for reasoning tasks. The main contribution is the use of Knowledge Representation Reasoning with LP theories modelling the knowledge of the user agent (beliefs, intentions, and expectations) to reason, plan and to optimally solve the DC problem. Another contribution is the design of a system component that extends the theory of mind, for the user model, with emotions to detect if the user decepts to the system or to itself. This component has the aim to alert and inform the facilitator when E-friend detects possible deceit signals from the student. E-friend was designed to help first year university students to manage stress/anxiety to optimal well-being development and attempt the prevention of depression and addictions leading. Students can interact through a chat-bot (text-based questions and answers) to help the system learns from the user, at the same time the user learns from itself improving mental health well-being.

Keywords: Agent Systems Architecture · Agent Knowledge Representation · Agent Reasoning Planning System · Dialogue Composition · Logic Programming · Human-Agent Interaction · Artificial Intelligence for detecting deception · Well-being / Mental Health Optimization · Emotion modeling

1 Introduction.

The research works of the World Health Organization (WHO) referred in [28] have concluded that stress is the world mental health disease of the 21st century and may be the trigger for depression and even suicide if it is not treated correctly. WHO estimates that, in the world, suicide is the second cause of death in the group of 15 to 29 years of age and that more than 800,000 people die due to suicide every year. Also that mental illnesses generate high economic losses since sick people and those who care for them reduce their productivity both at home and at work. According to data from the WHO, 450 million people in the world, suffer from at least one mental disorder. Well-being (meaning having no anxiety, depression or stress) and physical health have been studied

by many Scientists. For instance, Elizabeth H. Blackburn, Carol W. Greider and Jack W. Szostak (all Nobel prize) have shown that Telomerase activity is a predictor of long-term cellular viability, which decreases with chronic psychological distress [11]. E. H. Blackburn et. al. proved that mindfulness may exert effects on telomerase activity through variables involved in the stress appraisal process [11].

Over the years, science has shown that the brain and the mind work synergistically, that is why the brain can be reorganized, re-educated and regenerated by forming new nerve connections or paths when learning to control the mind through therapies. There are different successful techniques to support a student in overcoming their psychological difficulties such as referred in [18] [16] [17]: *Mindfulness* and *Cognitive Therapy*.

Our approach, called E-friend, proposes the use of Mindfulness and Cognitive Behaviour Therapy (CBT) to help the student to improve his/her mental health, see [18] [16]. It also includes techniques referred in [18] [16] for finding the Element, reaching Flow states, Silence in Therapy and Poetry Therapy provide value added to our proposal particularly for college students. Mindfulness and Flow States are independent different behaviours however they can be alternated [18][16]. Another element considered in E-friend is *Theory of Mind (ToM)* [23], which is the ability of humans to ascribe elements such as beliefs, desires, and intentions, and relations between these elements to other human agents. *Simulation Theory of Mind (ST)* [23] emphasizes the process of putting oneself into another's shoes'. ST argues for a simulation based method for model selection. E-friend has been designed to have a ToM of the User Agent (the student) as a Logic Programming (LP) Theory in the User Model. It is by logic theories in terms of logic programs that is possible to reason and plan DC to help the user human development deducing from her beliefs, intentions and desires what she is thinking for decision making. In this case, ST is modelled by reasoning using LP.

E-friend also has the capacity to answer questions of university matters and try to create a link with the student because it considers her pleasures and hobbies. Enriching talks ("mild Therapies"), proposed to be used by E-Friend, are mainly based on mindfulness + cognitive therapy and advice in the professional career preferences, which are focused particularly on preventing and managing mild symptoms of stress, anxiety, and depression to reduce the risk of failure in the university life due problems in learning, and to optimize mental health and behaviour when they face the university challenges as it is justified in [22]. Thus, during sessions with E-friend, it is intended that the students understand, accept and "become a friend of" their minds and emotions obtaining a better performance both in school and in their personal life. As referred in [13] in the work of Luksha Pavel et al these existential skills include an ability to set and achieve goals (willpower), self-awareness / self-reflection ability (mindfulness), an ability to learn / unlearn / relearn (self-development) relevant skills (e.g. skill-formation ability), and more. Well-being is a skill to be learned as explianed in [16] and this topic is explored in the context of a chatbot in [17]. The core type of dialogue for every dialogue session of E-friend is *Maieutics*. However each single agent task micro-dialogue as secondary type of dialogue can be one of the following according the categories stated by Douglas Walton referred in [27]: Persuasion. Inquiry, Discovery, Information-Seeking, Casual chat, Negotiation, Deliberation, Eristic.

E-friend has been deigned not for lying nor deceiving since the perception of the student of being betrayed by the system could bring negative psychological consequences and it is in opposition to the presented mild therapies [21]. Also deception to the student would be against the aim of our system to help the student to be compassionate for not manipulating the others but to take care of herself and the others. We believe that it can be possible to develop a system component to estimate when the student is lying to itself or the system and send an alarm to de facilitator. Whit this alarm the facilitator can be able to act in consequence helping the student to obtain better result during the session. This system component is totally independent to the existing E-friend software proving that E-friend's architecture is able to support updates without making important changes in the original structure.

Our paper is structured as follows: In section 2 we discuss chat-bots applied for mental health well-being, the use of emotions to complement a user model in AI systems and the Orthony, Clore, and Cllins (OCC) model [15] in logic programming for knowledge representation reasoning. In Section 3 we present the architecture of E-friend where x the Detection of User Deception Agent Module (DAM) is described. In Section 4 it is discussed how is extended the user model of theory of mind for handling emotions using the OCC model and Logic programming used for E-friend DC and DAM. Finally in section 5 we present our conclusions.

2 Related work

Applied Chatbots for Mental Health Well-being. As referred in [3] the work of Samuel Bell et al presents Woebot, a template-based chatbot delivering basic CBT, has demonstrated limited but positive clinical outcomes in students suffering from symptoms of depression. The work of Eileen Bendig et al referred in [4] presents promising areas for the use of chatbots in the psychotherapeutic context could be support for the prevention, treatment, and follow-up/relapse prevention of psychological problems and mental disorders. Also they could be used preventively in the future, for example for suicide prevention. According to the work of Samuel Bell et al in order to provide scalable treatment, several promising studies have demonstrated clinical efficacy of internet-based Cognitive Based Therapy, whereby the need for a face-to-face presence is negated. Woebot has demonstrated limited but positive clinical outcomes in students suffering from symptoms of depression.

In the work of Diano Federico et al referred in [5] it is presented an state of the art in mindfulness-based mobile applications and the design of a mindfulness mobile application to help emotional self-regulation in people suffering stressful situations. We invite the reader to check the work of Baskar Jayalakshmi et al referred in [1] where it is presented a comparison of Applied Agents implemented for improving mental health and well-being. In the work of Jingar Monika et al referred in [12] it is explored how an intelligent digital companion(agent) can support persons with stress-related exhaustion to manage daily activities. Also, it is explored how different individuals approach the task of designing their own tangible interfaces for communicating emotions with a digital companion. In the work of Inkster Becky referred in [10] it is presented an empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being

that is using mindfulness as mild therapy in combination with transfer to psychologist whenever the user ask for it. According to Samuel Bell et al several studies have investigated the clinical efficacy of remote-, internet- and chatbot-based therapy, but there are other factors, such as enjoyment and smoothness, that are important in a good therapy session.

Artificial Intelligent Applications complementing user profile with emotions. Emotions are considered high-level cognitive processes and lead an important part of human behaviour and reactions. These responses often are elicited when it needs rapid decision-making in our daily lives, for example, to be safe from danger or even to deceive. But also emotions are considered for long term decision making such as to buy the brand new cellphone or to get married. Emotions are useful to generate user profiles to have integral knowing about the person profiled. Meet emotional tendencies about a person for different situations could help to estimate if the person deceives based on behaviours opposed to the beliefs or knowledge of that person. Researchers use emotions in different areas, like in recommender system where authors studied the role of emotions and preferences in product appreciation [20]. The test and results show that considering emotions in the user profile improves the quality of product recommendation where emotions are fundamental in the act of purchasing. Emotions are personal characteristics, and for context-aware recommenders, emotions are modelled as contextual features, because emotions are dynamic and greatly influence user's preferences. This emotional feature has proven helpful in improving context-aware recommenders. In the last years, system developers have tried to improve human-computer interaction designing the systems around users (user-centred modelling). The idea here is to understand the user's behaviour under different conditions and scenarios for improving the user's experiences for a given system [14]. For this purpose, predictive models of user's reactions, responses, and in particular emotions, can help in the design of intuitive and usable systems. An emotional context is analysed and modelled based on varying types of sentiments that users express in their language. Analysing and modelling emotions, sentiments and personality traits provide valuable insight to improve the user experience in computer systems [14]. There is evidence of the need to adapt user interfaces by taking into account dynamic user features like emotions. It is crucial to go one step further in the adaptation of interfaces considering user's emotions at runtime. For this purpose, it is important to extend existing adaptation approaches to consider emotions. Like in [8], where the authors based their proposal in three main components: the inferring engine, the adaptation engine, and the interactive system. The authors consider three types of emotions: positive, negative, and neutral. The interface is able to identify if one particular adaptation has been found positive, negative or neutral by the user's feedback for future adaptations. These adaptations include widgets, fonts, colours, even the structure of the interface. The general adaptation process consists in using a kind of template base on contextual characteristics such as screen size or the brightness, using these principles, they compute appropriate adaptations, and then, they extend the interface considering emotions [8].

The OCC model, proposed by [15], is a psychological model frequently used in applications where emotional states can be simulated. The OCC model contains representations of the cognitive process where the emotional states represent the way of perceiv-

ing our environment [15]. This model was designed to be implemented in a computer language; this implementation is oriented, but not limited, to a rule-based system and provide examples of these rules [15]. For example, considering the emotions guilt and shame (these two emotions are felt by persons who lie [9]), the OCC model groups this two emotion into one emotion set called shame. The cognitive process to elicit shame is a valence evaluation of disapproval of self-censurable action that is expressed by one element (called token) in the set. This token depends on the emotion intensity and the OCC model provide three variables to compute that intensity: 1) censurability degree, 2) the cognitive strength, and 3) the deviations of the agent's action from expectations based on the person or social role [15]. In the same way, it is possible to model the emotions of fear and excitement identified in the lie detection process [26]. Model emotions computationally, it is possible to feed expert systems, intelligent agents, or robots to improve lie detection, and complete user's profiles. The work of [24] uses the OCC model for modelling emotions using LP under the Answer Set Programming (ASP) to generate stories with morals. However this work is not used for modelling a User Agent Profile or Mind Theory. The work of [23] presents Multi-Agents architecture based on theory of mind to model deception reasoning using logic. Literature of related work does not include reasoning logical based architecture for dialogue composition.

According to the above, it is possible to design a rule-based module to complete emotional student profile using the OCC Model. Also, this module could be able to estimate when a student trying to deceiving itself or the system.

3 Design of Architectural Framework for a Logical Based Intelligent Agent System Chat-bot

The general architecture design of our system is presented in this section. Similar E-friend architectures have been presented in [18] and [16]. The contribution in the present work is to extend the architecture with an additional independent but collaborative Detection of User Deception Agent Module (DAM). Another Contribution is the definition of User Agent model considering a Theory of the mind [23] extended with Emotions [15] and [24] using Logic Theories to provide our Master Agent Composer and DAM the reasoning skill based in a symbolic-based Representation of the User Agent. The user agent model presented in our architecture is domain independent and will be discussed deeply in the next section. Consider that our domain independent architecture it is instantiated for the Mental Health Domain for optimizing students (User Agent) well-being.

3.1 E-friend general approach

We assume (in the general case) that the mathematical model behind the system is a form of an ideal (collaborative) game. This means we have two players (E-friend and the human agent). There are moves/decisions (alternatives among which each player chooses), and payoffs (numerical representations of the players' preferences among the possible outcomes of the game). A move taken by E-friend is a script of five questions/activities proposed to the student. The student chooses to do (or not) the activities

and/or answer (or not) the questions given by E-friend. Sometimes, E-friend allows the student to ask (restricted or open) questions to E-friend. We also assume that the ultimate (ideal) goal of the game from the point of view of E-friend is to help the student to obtain a better knowledge about itself and to improve its mental state.

In addition, we assume that normally the student desires to improve his/her mental state and trust E-friend to help him/her to achieve this goal, the last is based on teacher-students relationship experiences. However, we understand that we can have exceptions, as we explain below. The game assumes interactive decisions, applicable whenever the actions of two or more decision makers jointly determine an outcome that affects them all. Strategic reasoning amounts to deciding how to act to achieve a desired objective, taking into account how the student will act and the fact that he/she will also reason strategically. In this regard, the basic ideal assumptions are the following:

i. Both players have consistent preferences and are instrumentally rational in the sense of invariably choosing an alternative that maximizes their individual payoffs, relative to their knowledge and beliefs at the time.
ii. The specification of the game and the players' preferences and rationality are common knowledge among the players (explained under Common Knowledge).
iii. It is assumed that an intelligent agent in E-friend takes the decision of the move to make. Then, a team of agent-slaves perform such movement.

Due to the nature of our game (on the context of positive Psychology) we need to make three considerations that could affect our ideal situation. *First*, we lack of determinate game-theoretic solutions, and psychological theories and empirical evidence are therefore required to discover and understand how people play them. *Second*, human decision makers have bounded rationality and are rarely blessed with full common knowledge. We assume that the student not always necessarily choose strategies that maximize its payoffs. However, E-friend follows to maximize it payoffs (in helping the student to improve its emotional state). In particular E-friend computes it corresponding payoff by a kind of knapsack problem [16]. It selects the "best" 5 activities/questions in a particular suitable order. This is the move that it selects (so to speak) and the decision is made by our intelligent agent using an extended knapsack problem [16]. *Third*, human decision makers have other-regarding preferences and sometimes do not even try to maximize their personal payoffs, without regard to the payoffs of others, and psychological theory and empirical research are therefore required to provide a realistic account of real-life strategic interaction. The actual move is performed by a particular agent slave that is able to coordinate the 5 slaves that are in charge of each of the 5 micro-dialogues, respectively

Deception [23] is defined as the intention of a deceptive agent (A), to make another interrogator agent (B), to believe something is true that A believes is false, with the aim of achieving an ulterior goal or desire. In E-friend context, deception is considered as another preference. E-friend knows that, hence it should be prepared with tactics to discover this situation, and yet to still be able to help the student. To discover such a situation, E-friends has a basic theory of the mind of the student. Such model is constructed based on the profile of the student thanks to the information that the student has provided it in the different sessions. We also assume that E-friend has an emotional model (built using the OCC Model) of the student in order to try to understand it better.

All together plus some default logical rules provided by the coach, E-friend is able to construct a basic theory of the mind of the student, mentioned before. It can be identify that Logic programming appears to be a sound mechanism to represent all the above knowledge. This knowledge is structured using different theories that share common knowledge of lower level. At the top of this theories we propose the use of a knapsack problem to represent the final decision of the move to make.

Emotions analysis may help to lie detection when they are strongly felt, the problem is most people do not feel strong emotions at the time of lying, because most of them lie everyday making difficult to use emotions in micro-expressions for detecting lies [26]. But it is possible to identify inconsistencies in the interaction with the student. The component proposed to extend E-friend, must be able to identify this inconsistencies and represent it in an emotional model. As it can be seen in last section, the OCC Model establish that emotions are the result of environment evaluations. This model provide us with a cognitive structure about how emotions are generated and felt by humans. These processes can be represented as logical rules in logic programming languages. When User Deception Agent Module (DAM) identify the possibility of deception from the student (the intensity of emotions felt when humans lie, like shame, crosses some threshold), DAM sends an alert to the facilitator. With this alert, the facilitator can take actions in order to help the student to avoid deception. To demonstrate the feasibility of the idea and show its practical potential we follow a Proof of Concept strategy as it is presented in [16]. Recall that a proof of concept is usually small and may not be complete.

3.2 A Master-Slave AI design

We propose a master-salve conceptual design following a centralized approach. Namely, we create hundreds of slaves (at least one thousand) such that each of them can perform a very concrete task. All the tasks correspond to interactions with the students. Each interaction are specified as atomic micro-dialogues. An example could be simple or complex task such as to teach the student how to try a meditation exercise. Each task performed by a slave-agent is programmed in the ***Basic Script/Resources Language (BSRL)***. This is a low level language (to describe automaton) invented for this purpose. Some of these slave-agents can be created by humans and some other can be constructed semi-automatically[5]. Associated to each slave we have its Semantic Knowledge. For instance, slave named $E3$ could correspond to an exercise of sound-mindfulness, that belongs to the set of mindfulness exercises. Furthermore, the system has an explicit logical rule saying that this type of exercise normally helps to reduces anxiety, and so on. All the semantic knowledge of each slave plus a general theory of interaction among them is written in LP Language. So, the LP theory corresponds to the ***Master-Agent Artificial Intelligent Composer (MAIC)*** that *plans* a sequence of few tasks (for a 10-15 minutes estimated session) that are performed by our slaves that are presented (coordinated) by a distinguished slave (a program interpreter of BSRL in Python) to the student. An analogy that we can make is the following. The LP agent is like a *master composer* of a symphony for a particular audience. The pianist is a particular slave that performs a

[5] We do not consider this issue in this paper.

specific task (playing the piano). The director corresponds to our distinguished *slave* that actually coordinate the rest of *slaves*. After the execution of the symphony, according to the feedback (applause, reviews, etc.), the composer hopefully learns how to create a better symphony. E-friend has a domain-independent generic architecture that is specialized to an application domain and to a component library through a declarative domain theory (ASP code). The domain-independent generic architecture is embedded in a Knowledge Base (KB) written in ASP. In our case the application domain corresponds to our basic slaves written in BSRL. The main concrete tasks of our intelligent agent described in [18][16].

3.3 The E-Friend Intelligent Agent System Chat-bot

The E-Friend is a *Reasoning Planning System* that consist in a cycle of 4 sequential processes-modules, Fig. 1 a), plus the extension proposed: **DAM**. Fig. 1 b)

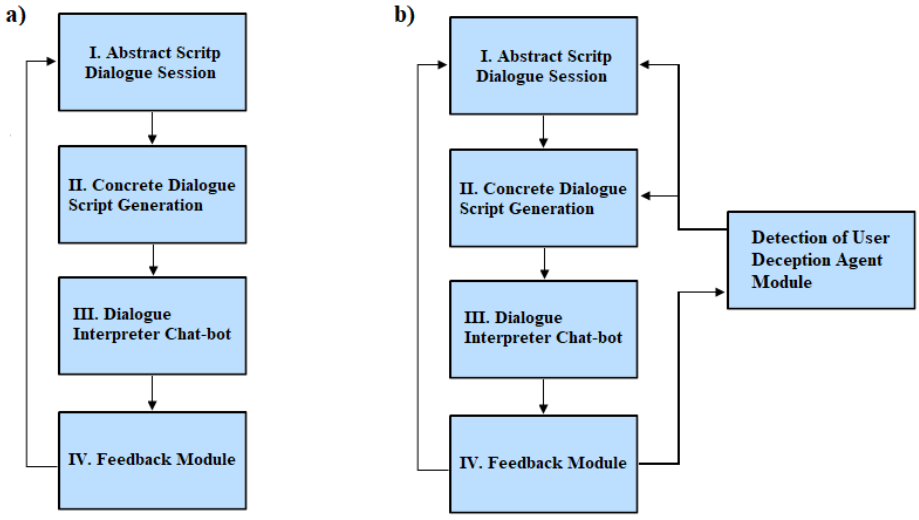

Fig. 1. a) The general architecture of the four sequential modules of E-friend. b) E-friend general architecture with the extension proposed (DAM).

I. Abstract Script Dialogue Session (ASDS) is generated by MAIC in this process. ASDS is a composition of slave agents tasks sequence to be performed by E-friend as a single dialogue session with the student. The name of each ***Slave Agent Interaction Task (AI-Task)*** (for example $E3$) corresponds to an interaction task with the student described in BSRL, the task is an interaction with the student (for example an enriching talk mindfulness exercise). Also each AI-Task is associated to semantic knowledge described in an ***Enriching Talks Theories Knowledge Base (ETalks-KB)*** in LP. Using the User Agent Model and the Enriching talks theories MAIC solves ***The Dialogue-Session Composition Problem (DCP)***. **Remark:** if E-friend detects a serious case of

anxiety-depression then it is proposed that the system transfers student to a human expert therapist (situation that is normally possible in a university institution). Even more the student could press an emergency button[6] available in E-friend that provides resources for getting immediate human experts help. The DCP is to compose a dialogue session as a sequence of AI-Tasks assets such that optimizes the contributions of the assets to the user with an optimal coherent, enjoyable and smoothable session. An optimal solution for DCP instance is the intended ASDS to be proposed by E-friend. ***II. Concrete Dialogue Script Generation*** translates the ASDS list to a single BRSL program. The idea behind the BSRL code (to describe automaton) is to define a basic programming language such that any program AI-task of our library is a highly malleable object where one could define and apply operators (mutation, crossover composition, selection, specialization, generalization). Briefly speaking each instruction in the BSRL code is a triple $< l, o, a >$ where l is a label, o an operator and a is an argument. BSRL resembles a kind of machine assembly language. ***III. Dialogue Interpreter Chat-bot*** corresponds to the director of the orchestra that executes the composed dialogue session as interactions of AI-Task with the student. ***IV. Feedback Module*** is an extraction process of relevant information and knowledge. This module filters a user conversation record to obtain the ***Student Profile State (SPS)*** . The SPS is used to provide feedback and to update the Extensional Knowledge Bases of the proposed System. The updates made in Extensional Knowledge Bases using the SPS could be about student profile, emotional status or history record, among others. This module also updates the answered questions Knowledge Base where also the questions made to the student are recorded. The Feedback Module is an extraction process of relevant information and knowledge. This module filters a user conversation record to obtain the **Student Profile State (SPS)** . The SPS is used for providing feedback and for updating the Extensional Knowledge Bases of the E-friend System. The updates made in Extensional Knowledge Bases using the SPS could be about student profile, emotional status or history record, among others. This module also updates the answered script questions Knowledge Base where the questions made to the student are recorded. ***The extension proposed: Detection of User Deception Agent Module (DAM)*** should be a special Agent Module that works as an observer of the dialogue between E-friend (modules 1-4) and the user Agent. This Agent module would be independent from the MAIC and the Slaves AI-tasks. It will be an intelligent Agent with the skill of reasoning for detecting User Agent deception. We conjecture that an extension of Theory of Mind[23] with a theory of emotions (the OCC Model) [15] and [24] would allow us to have a more refined model of the other User Agent (the student in our case) and with addition of detection of inconsistencies in answers of the User Agent to E-friend could help the E-friend Agent "coach" to diagnose / predict better when the student could try to deceive to the system about its progress regarding the support it is receiving from it. So by adopting the model in [23] and extending it with the OCC Model but also taking advantage of the benefits of ASP as presented in [24] we assume that it would enrich our system. It is important to highlight that the reference [23] is consistent with our approach because of its direct link to logic including the use of negation as failure (a

[6] A functionality learned from Woebot chat-bot presented in the work of Fitzpatrick Kathleen et al referred in [16]

key point in ASP). For instance we have modelled Theory of Mind, beliefs, desires, ignorance, trust, theory-theory, simulation theory, totum ex parte, to mention some of them using negation as failure architecture. The following is an example of encoding of the OCC Model extending theory of mind using beliefs and desires with emotions: $joy(Agent, Consequence, T) :- agent(Agent),\ belief(Agent, Consequence, T),$ $desire(Agent, Consequence, T),\ time(T)$. 'if Agent believes and desires a Consequence then the Agent will feel joy'. This module makes diagnoses by assigning values to the actions (questions-answers) of the user through time and the different sessions.The diagnosis when assessing the possibility that the user is deceiving uses 5 values of traffic light type (green, yellow - phases 1,2,3 - and red). For example, if the traffic light is yellow phase two, the observer module sends E-friend permission to ask the user questions and check if there is consistency between what he is saying and what he is doing or making. DAM deliberate/reason by itself using the theory of the mind of the user profile extended with emotions. When the DAM recognizes negative emotions on the part of the user while the student is interacting with the E-friend (for example in the attitude when the student is performing a mindfulness exercise), DAM does not make decisions towards the user but it sends a warning message to a human facilitator, in this case facilitator could or not, decides to take action on how to help the user in its human development. When DAM detects emotions that could conjecture the student is deceiving to the system or deceiving to himself regarding the achievements and benefits of practising a suggested exercise then the traffic light could go down to red, in this case, the warning becomes an alert so facilitator must take action. It is proposed to instantiate in this architecture the use of Answer Set Programming (ASP) [25] to provide Reasoning and Knowledge Representation in the ASP module. Logic rules in these theories are defined by psychologist while others are inferred from the student feedback (for example to evaluate the smoothness between two AI-tasks).

Page 26 of [7] presents a diagram of any addiction cycle based on different psychological theories. The cycle is a whirlpool, each turn becomes more aggressive even in that degenerative cycle until the end point is schizophrenia, death or prison. There are several inconsistencies where the person suffering addiction is deceiving to herself and can be helped by DAM to resolve them at different levels of the addictive cycle regarding to: a) Principles of caring for oneself and others. False thoughts caused by overprotective/over-demanding parenthood, b) Harmful guilt self-judgements and complexes (the image of a certain psychic situation which is strongly accentuated emotionally and is, moreover, incompatible with the habitual attitude of consciousness), and c) To believe that a addictive substance trigger will give peace and resolve inconsistencies in a) and b). In the cycle: 1) obsession with harmful judgements, 2) Imagination ritual considering the imaginary-emotive pleasure that consumption will give in various situations of the trigger in c), 3-4) to believe that pleasure gives personal development when, in fact, due to the disordered form of its consumption, it does not resolve a) and b), but it damages the body and disgust the body over time, making her look for a new more aggressive trigger, and 5) Emotional whirlwind: handling remorse, guilt, self-justification, despair, shame and promises based on a and b but also being inconsistent in finding the solution to the problem and therefore inconsistent in a coherence between beliefs and acting.

4 Answer Set Programming Solution Proposal

Student model. Having a user model is fundamental in this kind of personalized systems as we have learned and adjusted to our particular case from the presented in [2]. The model is presented in our architecture for the particular instantiated Mental Health Domain for students well-being. Below it will be referred the User Agent as student. The system assumes a student model that is almost empty at the beginning but is updated by the interaction between the student and the system. This model is constructed based on the preferences and needs of the student. Recall that the major point (and almost the unique) is to help the student to know herself better. See Table 1 for a concrete fictive student Alejandra. It is proposed to consider emotional modelling to complement the user profile and adapt the system responses based on the emotional states of the user. This could be useful at the moment to estimate the possibility that the user deceits the system. It can be obtained emotional information of the user through surveys, like in the above works, and complete profile and improve the quality of the system. Once the information is gathered, it can be modelled both emotions and cognitive processes that generate them computationally using the OCC Model [15]. Conceptually our ASP program uses a complex theory that involves Student Profile State, Knowledge of Enriching Talk theories (Mindfulness, CBT, 'Flow states', etc.) that will solve the DCP presented in Section 3. **Remark.** DCP optimization builds and propose a student dialogue session

Table 1. Student model generated by E-friend based on the preferences and needs of the student.

Student Model[7]**.**

Alias:	Ale
Gender:	Female
Sleep:	Important
'Finding Element':	High interest
Flow:	not interested
Mindfulness:	medium Interest
Level MFNS:	second week
Preferred MFNS exercise:	Love and compassion
Cognitive Therapy:	low interest
Hobby:	Movies
FLAG status:	Green
Concrete emotional status:	Improving

Activities to support.

1. Sleep proposed routines (listening to ocean wawes)
2. Casual conversation about movies
3. exercises "Finding Element"
4. Mindfulness information and exercises
5. Academic information and support (advices)

script. The recommendations are derived from the diagnosis of emotional type theory and treat on what activities and conversations to follow for achieving and maximizing a good emotional status. The logistic determined by logical rules will provide for a good

emotional status exercises, for a regular emotional status supportive exercises and for bad emotional status coaching assistance though chat-bot dialogue.

Enriching Talks Theories in Answer Sets Programming (ASP). It has been defined for this project five logic programming theories under ASP semantics to model the student profile and construct the chat script proposal for each session. In Table 2 examples of logic rules within E-friend are shown.

Hobbies Theory. Suggest exercises and conversation recommendations encompassing the student likes. **Build Script Theory.** Assemble activities suggested to the student, describes rules for assembling the AI-tasks to build the final script dialogue, logistic and partial order of dependency between activities that can be developed by the student through the advance of dialogue sessions. **Emotional Well-being Theory.** It states description in logic of the OCC Model of emotion [15] that has an objective to achieve and maximize happiness of the student. It is considered, inferred, modeled and represented the knowledge of intentions, beliefs and desires of the student. To achieve this, we implement an adaptation to the OCC Model [15] and [24], using answer set programming(ASP), and encode the student conversations sessions in terms of this model. **Diagnosis of Emotional Type Theory**. Diagnose the emotional status of the student inferred from the student conversation with the chat-bot. It keep track of the emotional status of the student through past conversation sessions to make a better diagnose. The description in logic defines three cases for emotional status: good, regular and bad. It is considered and processed answers given by the student during conversation sessions, the self appreciation reported by the student during conversation sessions and the participation response to the proposed dialogue during conversation sessions. **Well-being Theory.** Suggest mild therapies AI-task interactions to compose a enriching talk dialogue session considering the feedback of student profile. **Academic Theory.** Suggest AI-task interactions with the aim to upgrade the scholar status of the student that can be considered to compose a dialogue session considering academic and emotional student profile. **Empathy Theory.** Suggest AI-task interactions with the aim of strengthening the empathy with E-friend but mainly help the student to achieve a healthy emotional status. **Causal Chat Theory.** Suggest a AI-task interactions of casual chat dialogue within the student dialogue session having the main purpose of retrieving from the student relevant information that is out from the scope retrieval of traditional specific domain theories. **Prescription Theory**. It states description in logic of rules to infer and make recommendations that will be considered for building the student session script, the recommendations are derived from the diagnosis of emotional type theory and treat on what activities and conversations to follow for achieving and maximizing a good emotional status. The logistic determined by logical rules will provide for a good emotional status exercises, for a regular emotional status supportive exercises and for bad emotional status coaching assistance though chat-bot dialogue. Logical rules are examples of E-friend Enriching Talks Theories present in the ETalks-KB. Updates and believe revision are also fundamental concepts required in our application. There are many well known proposed solutions based on ASP such as [19]. There is promising work related to *ethical* chatbots [6] that could allow E-friend to become more respectful to improve its ethical interaction with the student.

Table 2. Example of logical rules for E-friend Enriching Talks Theories in ASP.

Logic Rule Example of Enriching Talks Theories in ASP
Well-being
E-friend normally proposes a mindfulness exercise of breathing due to observed high long term anxiety unless exception happens. `candidate(T) :- mindfulness(T), long_term_anxiety(high), belongs_category(T,Y), mindfulness_breath(Y), not exception_well_being(T).`
Hobbie
E-friend normally proposes a movie related to a gender of movies liked by the student that has not been watched yet unless exception happens. `candidate(T) :- movie(T), movie_category(Y), movie_example(T,Y), middle_may_like(Y), not movie_watched(T), not exception_hobbie(T).`
Diagnostic[a]
E-friend diagnoses a status of positive stress for the student due to observations of student tackling academic university workload and green global emotional status. `diagnostic(postive_stress) :- university_workload(light,difficult), completed_academic_assignments(success),student_perception(positive_stress) global_emotional_status(green).`
Academic
E-friend normally proposes academic information specialized in university interships due observed green global emotional status unless academic exception happens. `candidate(T) :- academic_info(T),belongs_category(T,university_interships), global_emotional_status(green), not exception_academic(T).`
Casual chat dialogue
E-friend normally proposes a casual chat dialogue of a new subject belonging a different category of skilled subjects talked during past sessions if her observed global emotional status is green unless exception happens. `candidate(T) :- new_subject(T),belongs_category(T,Y), last_subject(W), not belongs_category(W,Y), accomplished_skilled(W), global_emotional_status(green), not exception_casual_chat(T).`
Empathy
E-friend proposes animated gif resource for generating motivation empathy due the diagnostic for the student is anxiety but her global emotional status is green unless exception happens. `candidate(T) :- belongs_category(T,gif), diagnostic(anxiety), gif_motivation(T), global_emotional_status(green), not exception_empathy(T).`
Emotions
E-friend generates emotions associated to deceipt, the cognitive structure is represented as logic rules in order to identify when the student tries to deceipt the system or itself. For example, a student feels shame if he/she carries out an action considered cheating, regardless of success. `shame_student(Action) :- action(Action), negate_action(Action,Negation), happens_student(Action), idl(Negation).`

[a] a).Possible domain values are Depression, Anxiety, Positive Stress, Negative Stress
b).Possible domain values for `global_emotional_status` are `green_flag`, `yellow_flag`, `red_flag` determined by the tendency progress of the student though the past sessions.

The following is an example of the interaction between the student and E-friend. First, it is provided the context of the Ale's history. Then, the analysis of the dialogue, shown in Table 3, from the DAM's point of view.

"It is well known that Ale is very cooperative and try all the well-being exercises that E-friend suggests her to do. E-friend asked him at the beginning of the semester how does Ale feels about his courses. Ale replied that that he is comfortable with them as well as with his professors. However, Ale expresses her insecurity about her mathematics course. Ale always had problems with Math. E-friend asked her again the same question after two weeks of classes and Ale gives the same answer as before. However, she mention this time that the professor of mathematics is very helpful and patient with her. After two more weeks, Ale is about to have partial exams. One of those days at a session (just before the exam), she says that she has hope to pass the exam. However, she says she is slow at exams that is his main concern. Ale and E-friend chat soon after she took her exam. She said to E-friend that she has doubts about having passed the exam. The exam had 10 questions and she answered seven of them. She had time to check six of them and she is positive about them. The last one that she tried, she answered in a hurry and she is not sure about it. She needs to have seven correct answers to pass the exam. Hence, she is in the borderline. She also said that she does not care about the grade as long as she accredited the exam. Ale and E-friend chat again, when Ale got her results..."

Table 3. Dialogue between the student and E-friend

Dialogue
1.- **E-fiend:** Tell me John, do you have the result of your math exam?
2.- **Ale:** Well, yes, I got the result yesterday.
3.- **E-friend:** How was it?
4.- **Ale:** I only got one wrong answer.
5.- **E-friend:** That is very good. I know that you did a great effort. Congratulations!
6.- **Ale:** Thank you.
7.- **E-friend:** Do you feel to share anything else to me?
8.- **Ale:** Tonight, I am going with some of my friends to celebrate the end of the partial exams and that we did well.

Based on the last context, Ale answered seven questions, and is sure she has six correct of those seven. Following the dialogue, in line four in Table 3, E-friend suspects that Ale failed the exam. She answered seven questions and has one wrong. There is no way to get seven right. Ale did not explicitly tell E-friend if she passed or failed the exam. E-friend, through DAM module, suspects that Ale tries to persuade, even deceit, to E-friend to think she passed the exam, because there is only one wrong answer in the exam. Maybe she passed the exam or maybe not. If she did, her responses were ambiguous. If she failed, then Ale may feel ashamed of herself and people may think she is not very smart. In this case, DAM should identify the ambiguity and would generate the emotion of shame according to the Emotions set of rules presented in Table 2, where

the action is ”to fail the exam”. E-friend will use the result provided by DAM and will decide inform the facilitator: it is possible that the student is deceiving the system or itself, and create the warning or the alert, depending of the traffic light type.

5 Conclusions and future work

One contribution of this work[8] is to propose an application for a Reasoning Logical Based Intelligent Agent System Chat-bot for Dialogue Composition named E-friend [9]. It is proposed within the Architecture the use of Logic Programming (LP) to provide reasoning skill to E-friend. Our main contribution in this work is the use of Knowledge Representation Reasoning with LP theories modelling the knowledge of the user (beliefs, intentions, expectations, emotions as an extended version of theory of the mind using the OCC Model of emotions) to reason, plan and to optimally solve the DC problem. Another contribution is the design of an independent component that complements the theory of the mind, within E-friend, [23] with emotions in the user model LP theories [15] and [24] to try to detect the user's deception to the system. This last component with the aim to help the user to avoid deception for achieving human development and mental health through cognitive compassionate skills. E-friend is instantiated in Well-being Mental Health domain for optimal well-being development of first year university students. As future work it can be designed detect deception module in E-friend using LP theory under Epistemic ASP to model deception pre-conditions presented in section 3.1 [23]. Also it could be used Possibilistic ASP or Probabilistic ASP to model Aggregating parameters presented in section 3.3 [23].

References

1. Baskar, J., Janols, R., Guerrero, E., Nieves, J.C., Lindgren, H.: A multipurpose goal model for personalised digital coaching. In: Agents and Multi-Agent Systems for Health Care, pp. 94–116. Springer (2017)
2. Baskar, J., Lindgren, H.: Cognitive architecture of an agent for human-agent dialogues. In: International Conference on Practical Applications of Agents and Multi-Agent Systems. pp. 89–100. Springer (2014)
3. Bell, S., Wood, C., Sarkar, A.: Perceptions of chatbots in therapy. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. p. LBW1712 (2019)
4. Bendig, E., Erb, B., Schulze-Thuesing, L., Baumeister, H.: The next generation: chatbots in clinical psychology and psychotherapy to foster mental health–a scoping review. Verhaltenstherapie pp. 1–13 (2019)
5. Diano, F., Ferrata, F., Calabretta, R.: The development of a mindfulness-based mobile application to learn emotional self-regulation. In: PSYCHOBIT (2019)
6. Dyoub, A., Costantini, S., Lisi, F.A.: Towards ethical machines via logic programming **306**, 333–339 (2019)
7. Fuentes, M.: The broken trap: The problematic of Sexual Addiction. Ediciones del Verbo Encarnado (2015), https://books.google.com.mx/books?id=Su2zCgAAQBAJ

[8] We thank the support of Psychologist Andres Munguia Barcenas.

[9] The E-friend application is available in https://github.com/luis-angel-montiel-moreno/efriend .

8. Galindo, J.A., Dupuy-Chessa, S., Céret, É.: Toward a ui adaptation approach driven by user emotions (2017)
9. Greenberg, A.E., Smeets, P., Zhurakhovska, L.: Lying, guilt, and shame. Tech. rep., Mimeo (2014)
10. Inkster, B., Sarda, S., Subramanian, V.: An empathy-driven, conversational artificial intelligence agent (wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR mHealth and uHealth **6**(11), e12106 (2018)
11. Jacobs, T.L., Epel, E.S., Lin, J., Blackburn, E.H., Wolkowitz, O.M., Bridwell, D.A., Zanesco, A.P., Aichele, S.R., Sahdra, B.K., MacLean, K.A., et al.: Intensive meditation training, immune cell telomerase activity, and psychological mediators. Psychoneuroendocrinology **36**(5), 664–681 (2011)
12. Jingar, M., Lindgren, H.: Tangible communication of emotions with a digital companion for managing stress: An exploratory co-design study. In: Proceedings of the 7th International Conference on Human-Agent Interaction. pp. 28–36 (2019)
13. Luksha, P., Cubista, J., Laszlo, A., Popovich, M., Ninenko, I.: Educational ecosystems for societal transformation (2017)
14. Mostafa, M., Crick, T., Calderon, A.C., Oatley, G.: Incorporating emotion and personality-based analysis in user-centered modelling. In: Innovative techniques and applications of artificial intelligence. pp. 383–389. Springer (2016)
15. Ortony, A., Clore, G.L., Collins, A.: The cognitive structure of emotions. Cambridge university press (1990)
16. Osorio, M., Zepeda, C., Carballido, J.L.: Myubot: Towards an artificial intelligence agent system chat-bot for well-being and mental health. Accepted in Artificial Intelligence for HEaLth, PersonaLized MedIcine aNd WEllbeing Workshop at ECAI 2020. (2020)
17. Osorio, M., Zepeda, C., Carballido, J.L.: Towards a virtual companion system to give support during confinement. CONTIE (2020)
18. Osorio, M., Zepeda, C., Castillo, H., Cervantes, P., Carballido, J.L.: My university e-partner. In: CONTIE. pp. 150–1503 (2019)
19. Osorio, M., Cuevas, V.: Updates in answer set programming: An approach based on basic structural properties. Theory Pract. Log. Program. **7**(4), 451–479 (2007)
20. Poirson, E., Cunha, C.D.: A recommender approach based on customer emotions. Expert Syst. Appl. **122**, 281–288 (2019)
21. Rachman, S.: Betrayal: A psychological analysis. Behaviour Research and Therapy **48**(4), 304–311 (2010)
22. Ribeiro, I.J., Pereira, R., et al.: Stress and quality of life among university students: A systematic literature review. Health Professions Education **4**(2), 70–77 (2018)
23. Sarkadi, S., Panisson, A.R., Bordini, R.H., McBurney, P., Parsons, S., Chapman, M.: Modelling deception using theory of mind in multi-agent systems. AI Commun. **32**(4), 287–302 (2019)
24. Sarlej, M.: A lesson learned: Using emotions to generate stories with morals. Ph.D. thesis, PhD thesis, Computer Science & Engineering, University of New South Wales (2014)
25. Schaub, T., Woltran, S.: Answer set programming unleashed! KI **32**(2-3), 105–108 (2018)
26. Vrij, A.: Why professionals fail to catch liars and how they can improve. Legal and criminological psychology **9**(2), 159–181 (2004)
27. Walton, D.: The place of dialogue theory in logic, computer science and communication studies. Synthese **123**(3), 327–346 (2000)
28. World Health Organization: World health organization. WHO. https://www.who.int/ (2020)