

Aprentatge automàtic i agents conversacionals sensitius: un tutor virtual per a la sostenibilitat energètica

Dolça TELLOLS ASENSI (dtellols@gmail.com), Universitat de Barcelona

Resum Els *agents conversacionals sensitius* (“Sentient Embodied Conversational Agents” –SECA–), són personatges virtuals capaços d’establir converses complexes amb els usuaris i d’incorporar qualitats sensitives semblants a les dels humans. Aquest projecte formalitza l’arquitectura dels SECA i implementa una llibreria software que facilita la seva inclusió en aplicacions que puguin requerir comportaments proactius i sensitius per part de l’agent. Aprofitant l’arquitectura i la llibreria presentades, es dissenya, implementa i avalua un tutor virtual (la Terra) per fer més educativa una aplicació. Aquest projecte avalua un primer prototip de l’agent que utilitza tècniques de coincidència de patrons de paraules per analitzar el text dels usuaris i aprofita les dades recollides per millorar l’agent amb tècniques d’aprenentatge automàtic (“Machine Learning”). Els resultats de l’avaluació de l’aplicació amb el SECA Terra incorporat mostren com els participants estan satisfets tant pel que consideren que han après com per la seva experiència global.

1. Introducció

Els xat-bots són agents conversacionals no-personificats dissenyats per a comunicar-se amb l’usuari en interaccions simples i ràpides. Encara que poden ser útils, es queden curts quan es volen establir converses complexes amb els usuaris. Per a aquests casos, on cal centrar-se en la personificació i credibilitat dels agents, s’estan implementant agents conversacionals personificats [2] (“Embodied Conversational Agents” –ECA–), que són personatges virtuals capaços d’establir comunicació tant verbal com no-verbal amb l’usuari. No obstant això, els ECA solen estar creats ad hoc per a objectius específics sense opció a la seva reutilització o evolució dels seus components funcionals i estructurals.

Aquest treball estén una aplicació gamificada per a nens i nenes dedicada a recollir informació sobre els hàbits relacionats amb el consum energètic de famílies i a millorar la seva consciència energètica [7]. L’aplicació, anomenada “Bogeria energètica”, es va provar amb alumnes de diverses escoles [6] els quals manifestaren que els agradaria que expliqués conceptes que no entenien. Prenent-ho en consideració, aquest treball inclou a l’aplicació un tutor virtual que guia i instrueix els infants basant-se en un article anterior [1] que presenta l’estructura d’un ECA, enfocat a tasques, capaç de mostrar emocions i motivació. Aquest treball avança en l’estat de l’art dissenyant i implementant l’estructura d’uns ECA anomenats agents conversacionals sensitius (“Sentient Embodied Conversational Agents” –SECA–) capaços d’establir converses complexes estructurades, mostrar emocions, tenir necessitats i ser empàtics.

Els objectius principals d’aquest projecte són: proveir una arquitectura i una llibreria software per als SECA, i aprofitar-les per dissenyar i implementar un tutor virtual en el context de la sostenibilitat energètica. L’agent s’inclourà en una aplicació gamificada per fer l’experiència dels usuaris més educativa i per recollir més dades.

El treball segueix la metodologia del disseny centrat en l'usuari¹. En una primera iteració s'incorpora a l'aplicació un SECA que, per analitzar l'entrada de l'usuari, utilitza AIML (“Artificial Intelligence Markup Language”) i tècniques simples de detecció de patrons de paraules. Aquest prototip inicial permet recollir dades de converses per entrenar algoritmes d'aprenentatge automàtic (“Machine Learning”). Els models generats s'incorporen en una segona iteració del SECA per enriquir el seu processament del llenguatge natural (PLN). L'agent millorat també es revisa per avaluar l'experiència conversacional dels usuaris i l'eficàcia de l'arquitectura dels SECA. A més, es determina si la presència de l'agent afecta la quantitat de dades recollides.

2. Agents conversacionals sensitius

Definim un *agent conversacional sensitiu* (“Sentient Embodied Conversational Agent” –SECA–) com a un agent conversacional personificat (ECA) amb qualitats sensitives similars a les dels humans que li permeten percebre, “sentir” i respondre a certs factors.

Com a agent conversacional que és, un model d'arquitectura client-servidor s'adequa als requeriments de la plataforma SECA: modularitat, extensibilitat, visualització a temps real, rendiment i capacitat d'emmagatzematge. L'esquerra de la Fig. 1 mostra l'arquitectura dels SECA². L'aplicació del client els personifica en la interfície d'usuari i el servidor allotja els agents i té un controlador que comunica l'aplicació amb els SECA i les bases de dades. Addicionalment, la part dreta de la Fig. 1 detalla la llibreria dissenyada per implementar els SECA, la qual disposa de diversos mòduls i d'un controlador.

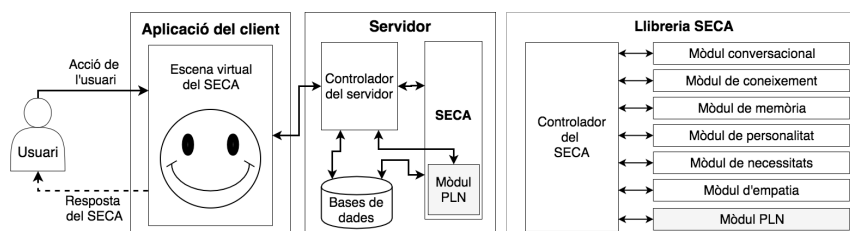


Figura 1. Esquerra: Arquitectura client-servidor. Dreta: Estructura de la llibreria SECA.

L'arquitectura i la llibreria proposades estan inspirades en METO [1], un ECA orientat a tasques, motivat i amb emocions. Aquest treball redissenya els mòduls de personalitat, necessitats i tasques (reanomenat com a mòdul conversacional), i crea els nous mòduls de coneixement, memòria, empatia, i processament del llenguatge natural.

El **mòdul conversacional** permet als SECA establir converses riques en llenguatge natural amb els usuaris. Aquest mòdul conté n converses juntament amb la referència a l'actual. Les converses es poden definir com un conjunt de *tipus de diàlegs (TD)*. Els TD poden ser proactius –si és l'agent el que els comença– o sota demanda de l'usuari, constitueixen un patró d'interacció especificat a través d'una màquina d'estats finita (MEF) i poden ser considerats com els blocs a utilitzar per construir converses dinàmiques: diverses converses poden reutilitzar els TD i aquests poden canviar durant una conversa. Per tant, les converses també s'especifiquen amb MEF (veure Fig. 2). Les transicions d'un

¹El disseny centrat en l'usuari (“User Centered Design”) és un procés de disseny centrat en les necessitats i els requeriments de l'usuari. És un procés iteratiu que consta de tres fases: anàlisi, prototipat i avaluació.

²Aquest treball implementa el servidor utilitzant el Framework de Python Django, i l'aplicació client amb Unity. REST (Representational State Transfer) API, basada en HTTP, s'encarrega de la comunicació. Una base de dades PostgreSQL guarda les dades de l'aplicació i una SQLite3 guarda informació de les interaccions del SECA, i el servidor utilitza llibreries Python com `scikit-learn`, `gensim`, `numpy` o `python-aiml`.

estat a un altre de les MEF són dinàmiques, depenen del context i simulen un flux de conversa natural gràcies a la implicació dels diversos mòduls del SECA.

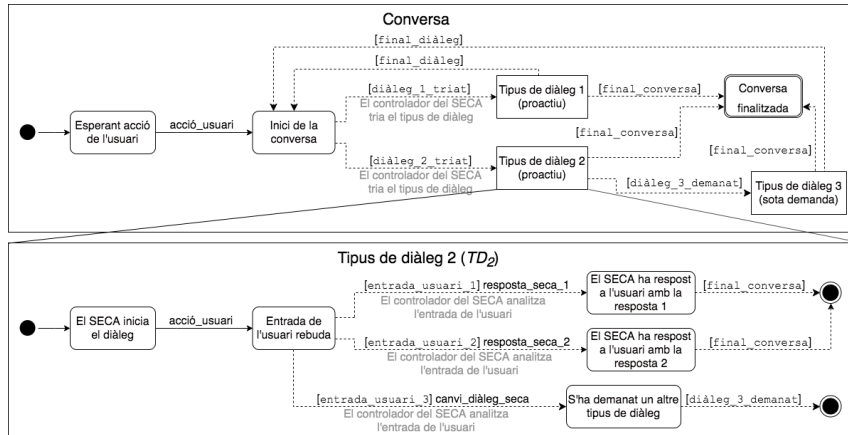


Figura 2. Esquema de diagrames de màquines d'estats finites d'una conversa i d'un TD.

El **mòdul de coneixement** gestiona el coneixement estàtic associat a les diverses converses. Conté una extensió d'AIML que aprofita algunes especificacions de METO AIML [1] per incloure paràmetres que poden tenir un impacte en els canvis en la personalitat de l'agent, i un conjunt de *components de coneixement* (C_j) predefïnits. Cada C_j és un conjunt de conceptes relacionats amb un cert tema.

El **mòdul de memòria** s'encarrega d'emmagatzemar informació dinàmica per evitar converses repetitives. En particular, gestiona les memòries a llarg i curt termini (*MLT* i *MCT* respectivament). L'*MLT* és un conjunt de *memòries* que es defineixen com una tupla formada per: un *id*; una *classe* que identifica si l'*id* correspon a un *TD* o a un concepte; i un comptador que s'inicialitza a 0. La funció *MCT* serveix per obtenir l'*id* que s'ha utilitzat més recentment d'una determinada classe.

El **mòdul de personalitat** proporciona trets de personalitat al SECA. La idea segueix el treball de Kshirsagar et al. [4] en definir personalitat, estats d'ànim i emocions com tres capes independents i està basada en METO [1]. Aquest treball l'estén de manera que els SECA tenen una personalitat fixada que permet un nombre diferent d'estats d'ànim i d'emocions que canvien gràcies a diverses matrius de transició probabilístiques.

El **mòdul de necessitats** gestiona les diverses necessitats que els agents puguin tenir i es basa en la jerarquia de necessitats de Maslow [5]. El mòdul té un conjunt de *necessitats* definides amb: una *etiqueta*; un *estat* que indica si la necessitat s'ha assolit (*inactiva*) o no (*activa*); i un comptador de *temps* que es reinicia al reactivar-se. Per exemple, el SECA podria tenir una necessitat d'"atenció" que s'activés al cap de 10 segons de no rebre interacció.

El **mòdul d'empatia** intenta endevinar els pensaments i sentiments dels usuaris a través de les seves interaccions. Igual que les necessitats, l'empatia es considera amb l'objectiu d'establir un vincle més fort amb l'usuari [3]. El mòdul disposa d'un conjunt d'indicadors d'"estat de l'usuari" els valors numèrics dels quals estan limitats i es controlen durant la interacció. Per exemple, es podria considerar el "cansament" dels participants, l'indicador del qual augmentaria a mesura que avancen les converses i podria utilitzar-se per precipitar-ne l'acabament si es fan excessivament llargues. Aquest treball assumeix que el màxim valor dels indicadors pot variar depenent de l'usuari.

El mòdul de processament del llenguatge natural (PLN) conté funcions dedicades a processar i a analitzar l'entrada de l'usuari.

3. Aprenentatge automàtic i processament del llenguatge natural

El mòdul de PLN de la llibreria SECA proveeix mètodes per carregar models i funcions que faciliten la implementació d'agents que entenguin diferents idiomes. Aquest projecte en concret se centra en la comprensió del català.

Primer es preprocessa l'entrada de l'usuari (el mòdul conversacional defineix converses estructurades predefinides que esperen una determinada entrada en certs estats). En una segona etapa, el problema es "redueix" a resoldre la classificació del missatge rebut perquè l'agent pugui donar la resposta més apropiada (el mòdul de coneixement conté les respostes predefinides que pot donar el SECA).

El preprocessament és necessari per facilitar la manipulació i l'anàlisi del text introduït per l'usuari, que es neteja i se separa en paraules ("tokenization"). Per a la posterior classificació, s'associa un vector de valors numèrics a l'entrada de l'usuari ("embedding"). Aquest, s'obté de la mitjana dels vectors obtinguts per cadascuna de les paraules amb *Word2Vec*³. Si *Word2Vec* no pot trobar l'"embedding" d'una paraula, s'aplica un senzill corrector ortogràfic sobre aquesta i es reintenta.

Obtingut l'"embedding" de l'entrada de l'usuari i el llistat de paraules preprocessades, per a la classificació, aquest treball proposa un sistema híbrid que utilitza funcions de cerca de patrons de paraules i tècniques d'aprenentatge automàtic. D'aquesta manera, depenent del problema a resoldre, s'intenta utilitzar la tècnica que funcioni millor. Utilitzant un mètode de cerca exhaustiu per triar els paràmetres més adequats en cada cas, s'han entrenat diferents algoritmes d'aprenentatge automàtic amb les dades disponibles. Encara que el sistema pot utilitzar-se en general, en aquest projecte s'han utilitzat dades de text en català per entrenar algoritmes de "Support Vector Machines", "Random Forests", "Gradient Tree Boosting", "Logistic Regression" i "Multi-Layer Perceptron". No s'han considerat tècniques com "Deep Learning" per la petita quantitat de dades disponibles. Els resultats dels tests s'avaluen en termes d'exactitud, precisió i exhaustivitat.

4. Tutor virtual per a una aplicació educativa en català

Aquest treball integra un tutor virtual en una aplicació per a infants⁴, amb els objectius principals de millorar l'experiència dels usuaris fent-la més educativa i d'avaluar l'arquitectura dels SECA. Des del panell principal de l'aplicació "Bogeria energètica" (vegeu esquerra de la Fig. 3), els infants poden interactuar amb el SECA Terra en tot moment.



Figura 3. Esquerra: Panell principal. Dreta: Conversació Terra-Usuari.

³Word2Vec [8] és una tècnica desenvolupada per T. Mikolov et. al que troba les representacions vectorials de paraules de bases de dades molt grans en un període de temps curt.

⁴Demostració en vídeo disponible en l'enllaç següent: <https://youtu.be/z6otygtACTo>

El disseny del **mòdul conversacional** de la Terra la fa proactiva en la majoria dels casos i motiva a reflexionar sobre temes relacionats amb l'energia. Incorpora 6 converses i 13 tipus de diàleg (*TD*). Alguns *TD* van destinats a ajudar a complir les tasques de l'aplicació i d'altres, per exemple, comparteixen recomanacions sobre l'eficiència energètica o revisen conceptes amb els usuaris. A més, hi ha *TD* sota demanda perquè l'usuari pugui preguntar directament a la Terra dubtes concrets. El **mòdul de coneixement** inclou *coneixement* de temes que agrupen conceptes i pràctiques relacionades amb l'energia i l'aplicació, i el de **memòria** conté les *memòries* dels conceptes i *TD* definits.

El **mòdul de personalitat** considera 3 estats d'ànim i 3 emocions, i les matrius de transició fan que la Terra se centri més en estats d'ànim extrems (feliç o trista) per fer més aparent l'impacte dels seus canvis afectius en l'experiència de l'usuari. El **mòdul de necessitats** gestiona la necessitat d'"atenció" de la Terra de socialitzar amb l'usuari i la necessitat d'"acabar el joc" que s'associa al desig d'ajudar l'usuari a complir els objectius de l'aplicació, i el d'**empatia**, per una part considera la "motivació" dels usuaris per controlar la freqüència d'aparició de la Terra durant l'aplicació i per una altra, guarda el "cansament" dels participants per precipitar l'acabament de les converses si es tornen excessivament llargues.

El **mòdul de PLN** considera 7 problemes de classificació diferents com ara l'"anàlisi de sentiments" o la "detecció d'explicacions". Per entrenar els models adequats, s'han fet servir 2285 missatges d'usuaris recollits amb el primer prototip de l'aplicació i s'han fet un total de 26640 experiments per trobar els millors paràmetres amb una cerca exhaustiva. Els algorismes de "Support Vector Machines" i "Multi-Layer Perceptron" són els que han donat millors resultats, aconseguint entre un 68.00% i un 95.27% d'exactitud dependent del problema de classificació i de les dades disponibles.

5. Avaluació

Aquest projecte ha testejat amb usuaris dos prototips de l'aplicació amb el SECA Terra incrustat. La Taula 1 il·lustra els objectius principals dels tests realitzats.

Taula 1. Resum dels diversos tests d'avaluació.

Test	Objectius	Participants
Pilot	Definir qüestionaris i comprovar-ne la concurrència	4
V1: Patrons de paraules	Recollir missatges dels usuaris per millorar el PLN	30
V2: PLN amb aprenentatge	Avaluació final (PLN i satisfacció dels usuaris)	15

Els participants, d'entre 10 i 12 anys, parlaven català, hi han participat voluntàriament i els seus pares han signat un consentiment on se'ls informava sobre el tractament anònim de les dades. Durant l'avaluació s'han obtingut dades qualitatives i quantitatives gràcies a dos qüestionaris que els participants responien abans i després de l'ús de l'aplicació, i a les dades recollides durant el període màxim de 8 dies d'ús.

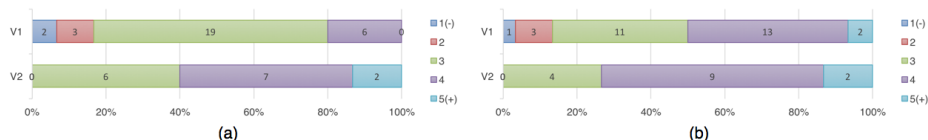


Figura 4. (a) Percepció de l'enteniment de la Terra. (b) Percepció de l'aprenentatge guanyat.

La Figura 4 (a) il·lustra com la percepció dels usuaris de l'enteniment de la Terra millora clarament en la versió amb la Terra que té el mòdul de PLN enriquit amb tècniques

d'aprenentatge automàtic (els usuaris passen de donar un 83.33% de puntuacions iguals o majors a 3 sobre 5, a un 100.00%). Per una altra banda, els resultats de la Fig. 4 (b) corroboren que els infants consideren que han après parlant amb la Terra (un 86.67%, en la V1, i un 100%, en la V2, han donat puntuacions iguals o superiors a 3).

A més, dels resultats anteriors, gràcies al qüestionari posterior, aquest treball comprova com la percepció dels participants respecte a l'experiència global de parlar amb la Terra ha estat considerada satisfactòria pel 93.33% dels participants, que l'han cridada voluntàriament una mitjana de 17 vegades, algunes amb la necessitat d'"atenció" activa.

Per últim, destacar que els usuaris han intercanviat amb la Terra un total de 7996 missatges i que ha augmentat la quantitat de dades recollides amb l'aplicació.

6. Conclusions i treball futur

Aquest article introdueix els SECA com a personatges virtuals capaços d'establir converses complexes i d'incorporar qualitats sensibles semblants a les dels humans. El treball presenta la seva arquitectura i proporciona una llibreria que facilita la seva inclusió en aplicacions que requereixen comportaments proactius i sensitius per part de l'agent.

El text il·lustra la proposta dissenyant, implementant i incorporant un tutor virtual (la Terra) en una aplicació educativa. Primer, avalua una versió de la Terra que utilitza només tècniques de coincidència de patrons de paraules per analitzar el text introduït pels usuaris i recull dades per entrenar algoritmes d'aprenentatge automàtic per a una versió millorada de la llibreria SECA amb el mòdul de PLN enriquit. Els resultats d'avaluar la nova versió mostren la satisfacció dels participants i verifiquen com s'ha enriquit l'experiència de l'usuari donat que la majoria consideren que han après interactuant amb la Terra. En definitiva, es confirma l'eficàcia del disseny i que pot ser positiu atorgar qualitats semblants a les dels humans com personalitat, necessitats o empatia per augmentar el vincle amb l'usuari.

Com a treball futur es podrien enriquir mòduls com el conversacional a través de l'addició de noves converses i *TD* o el de personalitat per considerar agents amb personalitats diferents depenent de l'usuari. Pel que fa al processament del llenguatge natural, presenta moltes dificultats i, encara que el nou sistema ha reduït la ràtio d'error de les respostes del 22.31% a l'11.46%, encara hi ha opció de millorar amb altres tècniques.

Referències

- [1] P. Almajano, D. Tellols, I. Rodríguez, and M. Lopez-Sanchez. *Meto: A motivated and emotional task-oriented 3d agent*. In *Recent Advances in Artificial Intelligence Research and Development*, volume 300, pages 263–268. IOS Press, 2017.
- [2] J. Cassell. Embodied conversational agents: representation and intelligence in user interfaces. *AI magazine*, 22(4):67, 2001.
- [3] N. Feshbach and S. Feshbach. Empathy and education. *The social neuroscience of empathy*, 85:98, 2009.
- [4] S. Kshirsagar. A multilayer personality model. In *Proceedings of the 2nd international symposium on Smart graphics*, pages 107–115. ACM, 2002.
- [5] A. H Maslow. A theory of human motivation. *Psychological review*, 50(4):370, 1943.
- [6] K. Samsó. Design, implementation and evaluation of a gamified digital cultural probe in the context of energy consumption. 2017.
- [7] K. Samsó, I. Rodríguez, A. Puig, D. Tellols, F. Escribano, and S. Alloza. From cultural probes tasks to gamified virtual energy missions. In *Proc. British Computer Society HCI Conf.*, page 79, 2017.
- [8] M. Tomas, C. Kai, C. Greg, and D. Jeffrey. Efficient estimation of word representations in vector space. *CoRR*, abs/1301.3781, 2013.