

ÈTICA A LA ROBÒTICA



Víctor López Bernal
Diego Saldaña Taipe
Tutora: Encarna García
17-12-2019
Batxillerat Tecnològic

Agraïments

Per començar, agraïm a la nostra tutora Encarna García pel suport i la seva contribució en el desenvolupament del projecte, així com els seus consells i ensenyaments respecte a determinats aspectes de la part pràctica i de redacció.

També agraïm a la nostra família per seguir de prop el nostre treball i en ajudar-nos a l'hora de practicar l'exposició i donar idees i consells quan els necessitàvem.

Tots aquells que van respondre el formulari que vam fer tenen el nostre agraïment, ja que sense les seves respostes no hauríem pogut elaborar la part més important del projecte.

Finalment, agraïm al MIT per inspirar-nos en la direcció que havia de portar el treball i ser un dels primers en endinsar-se en el futurista àmbit de la robòtica. Els autors d'Arduino també han jugat un paper important per a saber com programar i fabricar la part física del projecte i gràcies per últim a l'escriptor Isaac Asimov, ja que sense ell no coneixeríem les bases ètiques de les màquines com ho fem avui dia.

Resum

El treball tracta sobre l'ètica aplicada a la Robòtica en la societat, els primers principis ètics que van ser formulats respecte les màquines, les accions que han pres els països per adaptar-se a l'evolució tecnològica, arrel els problemes i riscos que poden sorgir, i en la creació d'una intel·ligència artificial (IA) que representi la culminació dels marcs jurídics que haurien d'estar aplicats als robots autònoms.

Es consulta el treball de l'Institut Tecnològic de Massachusetts (MIT): *Moral Machine*, sobre la creació d'una intel·ligència artificial on el seu criteri estigui format per les decisions de la població, recollint les respostes a través d'una pàgina en forma de qüestionari i classificant-les respecte els factors demogràfics, geogràfics i culturals.

Seguint l'exemple s'elabora un prototip d'intel·ligència artificial (IA) anomenada *Autem*, on el seu criteri d'elecció estigui format per les decisions de la població que va contestar un qüestionari on s'exposaven situacions ètiques. Arrel les respostes s'analitzen les raons de perquè la intel·ligència artificial hauria d'actuar d'una forma determinada i seguidament es passa al seu disseny, implementació i programació utilitzant una tarja Arduino UNO.

Finalment, es fa una conclusió de cada apartat, fent una reflexió sobre allò tractat, i es sintetitza la programació final d'*Autem* i les seves bases ètiques.

Paraules clau: ètica, Robòtica, intel·ligència artificial (IA), Institut Tecnològic de Massachusetts (MIT), *Moral Machine*, Arduino.

Abstract

The project is about the ethic applied to Robotics in our society, the first ethic principles that were expressed about the machines, the actions that the countries have made to adapt to the technological evolution, following the problems and risks that could arise, and in the creation of an artificial intelligence (AI) that represents the culmination of the juridical frames that would have to be applied on the autonomous robots.

The *Moral Machine*, the work of the Massachusetts Institute of Technology (MIT) is consulted. This work is about the creation of an artificial intelligence whose criteria was made from the decisions of the population, which were collected through a page with a questionnaire, and classifying them according to demographic, geographic and cultural factors.

Following the example, a prototype of an artificial intelligence is elaborated. It is called *Autem*, and its choosing criteria is formed by the decisions of the population that answered a questionnaire where ethic situations were exposed. From the answers an analysis is made about the reasons why the artificial intelligence should act in a determinate way and then it goes to the design, both the physical and the programming part, using an Arduino UNO board.

Finally, a conclusion is made of each section, making a reflection about the things that have been worked on, and the final program and the ethical bases of *Autem* are synthetized.

Key words: ethic, Robotics, artificial intelligence (AI), Massachusetts Institute of Technology (MIT), *Moral Machine*, Arduino.

Índex

Introducció	3
1. Les Lleis de la Robòtica	5
1.1. Les Lleis de la Robòtica d'Isaac Asimov	6
1.2. Legislació sobre Robòtica	6
1.2.1. Regne Unit i la robòtica	7
1.2.2. Legislació sobre robòtica a la Unió Europea	7
1.2.3. Legislació sobre robòtica a EEUU	9
2. El projecte <i>Moral Machine</i>	10
2.1. Els escenaris	10
2.2. Anàlisi dels resultats	13
2.3. Anàlisi en funció del tipus de població enquestada	15
3. AI: Autem	18
3.1. Medicina i sanitat	18
3.1.1. Eutanàsia: Petició de mort digna a un robot infermer	19
3.1.2. Part: Complicacions letals com a robot comadrona	19
3.2. Feines de risc:	20
3.2.1. Construcció: Accident de caiguda de pes davant un robot obrer	21
3.2.2. Mineria: Extracció de minerals a Mart amb un robot miner	21
3.2.3. Salvament en entorns radioactius: Risc de catàstrofe nuclear amb robot funcionari	22
3.3. Militar i seguretat:	23
3.3.1. Bombardeig: Atac d'objectius amb població civil	23
3.3.2. Seguretat: Situació d'ostatge amb robot de seguretat	24
3.3.3 Presó: Reducció de revolta com a robot funcionari	24
3.4. Anàlisi dels resultats de l'enquesta	25
3.4.1. Àmbit 1: Medicina i Sanitat:	26

3.4.2. Àmbit 2: Feines de risc	27
3.4.3. Àmbit 3: Militar i Seguretat	29
3.5. Les Lleis d'Autem	31
4.- Disseny d'Autem	33
4.1. Plataforma de desenvolupament de projectes Arduino	33
4.1.1. Descripció de la placa	34
4.2. Arxius d'àudio	34
4.3. Disseny del circuit	36
4.5. Programació	37
4.5.1. Programació en C:	37
4.5.2. IDE Arduino:	38
Conclusions	39
Propostes de millora	40
Referències bibliogràfiques	41
Annex 1: Formulari Ètica a la Robòtica	
Annex 2: Respostes del formulari	
Annex 3: Guió d'Autem	
Annex 4: Mapa conceptual de programació	
Annex 5: Programació d'Autem	

Introducció

La robòtica és i serà part de la nostra societat, prendrà protagonisme en diversos aspectes de les nostres vides i el seu desenvolupament haurà de ser responsable, especialment el de la seva programació: els sistemes d'intel·ligència artificial (IA).

Aquest tipus de sistemes seran explotats i molt utilitzats per la seva capacitat d'aprenentatge i autonomia a les accions. No caldrà escriure específicament què haurà de fer la màquina, és a dir, tindrà cert criteri propi.

És per això que vam considerar vital analitzar detingudament aquest sector de la tecnologia ara que es troba en una vertiginosa expansió.

Al tenir els nostres plans de futur envers la tecnologia, nosaltres, Víctor i Diego, vam decidir desenvolupar un projecte de recerca que no només pertanyés a l'àmbit tecnològic sinó que també impliqués treballar sobre qüestions actuals, que no haguessin estat molt definides i que aprenguéssim més sobre la programació i la robòtica.

Com ja ha estat mencionat, la manca d'anàlisi d'aquest àmbit fou una de les raons per les que es va iniciar el projecte. També ens semblà molt interessant comprovar com es podia desenvolupar la filosofia ètica en un "entorn" on la mecanització i la tecnologia són els elements predominants.

Específicament, els nostres objectius són:

- Analitzar l'ètica implantada a la robòtica, des dels seus orígens fins al dia d'avui, estudiar les mesures proposades per diferents països i seguir el seu desenvolupament.
- Elaborar la nostra pròpia "intel·ligència artificial", després de recopilar i estudiar la informació necessària, representant la culminació de tot el nostre treball com a prova vigent dels nostres esforços i conclusions.
- Aprendre més sobre la programació i la creació de circuits, així com familiaritzar-nos amb els programes que utilitzéssim.

L'estructura del nostre treball es fonamenta en els nostres objectius i es desenvoluparà en quatre parts diferenciades. En primer lloc, es realitzarà l'anàlisi de la legislació actual que es fonamenta a Les Lleis de la robòtica enunciades per Isaac Asimov a la seva novel·la *Jo Robot*. A continuació, exposarem els fonaments teòrics del nostre treball i l'exemple que vam seguir del treball de l'Institut Tecnològic de Massachusetts: la *Moral Machine*. Després passarem al desenvolupament teòric de la nostra pròpia intel·ligència artificial, establint com hauria de ser la seva base ètica en base a una enquesta difosa entre els usuaris. Finalment, ens dedicarem al disseny de la part física de la nostra IA i a la redacció de la seva programació.

1. Les Lleis de la Robòtica

Com ja s'ha mencionat abans, la robòtica avança cada vegada més de pressa. Les màquines comencen a substituir als humans en oficis i continuarà així en els anys venidors.

Són ben variats els camps on aquestes màquines poden superar-nos, el més polèmic i on la gent es mostra més preocupada és en l'àmbit militar.

Fa poc que robots dirigits a distància per persones són utilitzats en conflictes bèl·lics, on la seva especialitat és l'eliminació d'éssers humans. Encara no ser totalment autònoms, la intel·ligència artificial o AI està també en desenvolupament i en diversos casos ha estat implantada en dispositius tecnològics (com la Siri dels dispositius Apple).

A causa d'això, el fet de que els robots obtinguin una autonomia basada en una programació predeterminada fa desconfiar a un gran nombre de persones: com un robot pot saber quina és la millor opció en un dilema ètic? Si un tren viatgés per un carril i s'hagués de prendre la decisió de passar per una via on dues persones resultarien mortes o per una altra on només hi hauria una víctima, si ja seria una decisió en certa manera difícil, com podria obrar un robot?

Moltes figures eminents han constatat que l'estudi d'aquest aspecte ha de ser crucial i previ a la creació de màquines per a camps anteriorment mencionats. Una forta protecció als valors morals hauria de ser implantada en la programació inicial dels autòmats.

Aquest raonament va ser plantejat fa setanta anys per l'escriptor d'origen rus Isaac Asimov, autor de l'obra *Jo Robot*, la qual ha resultat ser un dels primers recursos bibliogràfics del nostre treball.

Resumidament, la novel·la tracta sobre diferents situacions que desemboquen en dilemes ètics on persones i robots interaccionen entre ells per a superar-les. Una de les coses que totes les històries tenen en comú és que giren entorn a l'aspecte immutable a la programació dels robots, les anomenades Lleis de la Robòtica.

1.1. Les Lleis de la Robòtica d'Isaac Asimov

Les lleis de la Robòtica, mencionades a les novel·les de l'autor, han resultat l'epicentre del desenvolupament dels sistemes autònoms arrel la necessitat de crear certs marcs ètics.

Va ser a la seva novel·la *Jo Robot*, publicada per primera vegada el 1950, on es mostren les Lleis de la Robòtica:

1. Un robot no pot fer mal a un ser humà o, per inacció, permetre que un ser humà sofreixi dany.
2. Un robot ha d'obeir les ordres donades pels éssers humans, excepte si aquestes entressin en conflicte amb la primera Llei.
3. Un robot ha de protegir la seva pròpia existència en la mesura en que aquesta protecció no entri en conflicte amb la primera o segona Llei.

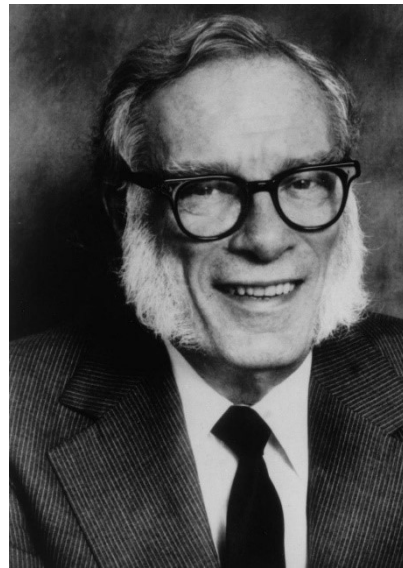


Figura 1.1. Fotografia de l'autor Isaac Asimov

1.2. Legislació sobre Robòtica

Durant els darrers anys, degut al ràpid desenvolupament que s'està produint en el camp de l'electrònica, la robòtica i la intel·ligència artificial ha sorgit la necessitat de regular jurídicament tots aquells temes relacionats èticament amb la utilització dels robots.

Sabent això, hem estudiat els intents d'acció per part de la societat actual, on destaquem a la Unió Europea, que ja va començar a debatre sobre escriure uns fonaments; Regne Unit, que va actuar de manera similar i els Estats Units, que principalment començaren a teoritzar sobre el tema.

1.2.1. Regne Unit i la robòtica

“L'ésser humà ha de prosperar”. Aquest és el fonament que la Real Societat i l'Acadèmia Britànica consideren necessari per a governar a les màquines intel·ligents dintre de poc presents en la nostra societat.

Es va redactar un informe en el qual es parlava sobre la possibilitat de la creació d'un nou organisme el qual pugui crear un marc ètic per al desenvolupament de tecnologies que utilitzin la intel·ligència artificial. El seu personal serien especialistes en la matèria.

El principal objectiu és assegurar una superioritat de les persones sobre els robots, eliminant qualsevol possibilitat de perill, ja que encara poder proveir molts beneficis, la robòtica obre l'existència de diversos possibles problemes.

L'aspecte comú que comparteixen els tres punts és que van partir de la idea de les Lleis de la Robòtica d'Isaac Asimov com a precedent i punt de partida per a la redacció de marcs ètics.

1.2.2. Legislació sobre robòtica a la Unió Europea

D'igual manera que amb Regne Unit, la comissió d'Affers Jurídics del Parlament Europeu va considerar important que la Unió Europea comencés a desenvolupar unes normes de seguretat, ètiques i legals acord amb els avenços que s'aconsegueixin en els camps de la robòtica i la intel·ligència artificial.

També es demana la possibilitat de poder aplicar la jurisdicció als robots més avançats en cas d'una situació de responsabilitat de danys.

Amb aquest argument, es van redactar un total de nou lleis en el document *European civil laws rules in robots*:

- 1. Protegir els éssers humans dels danys causats pels robots.** Aquest està basada en la dignitat humana com a fonament de tots els drets.
- 2. Respectar el rebuig de l'atenció per part d'un robot.** Aquesta llei estableix el dret d'una persona a negar-se a ser cuidada per un robot. Aquest tipus de casos sorgeixen arrel la presència del robot i incomoditat

provocada a la persona. Aquesta llei també afecta a aquells autòmats encarregats d'alimentar a ancians i discapacitats.

3. **Protegir la llibertat humana enfront els robots.** En situacions on el robot traspasa la llibertat d'una persona amb el pretext de protegir-la encara està en debat. La màquina ha de respectar l'autonomia de la persona, però el missatge ha d'estar adaptat en circumstàncies on hi ha un risc (com un alcohòlic), l'edat de la persona i el nivell de consciència i comprensió.
4. **Protegir la humanitat contra les violacions de la privacitat comeses per un robot.** Es creu important mantenir un límit en la intimitat d'una persona la qual es troba sota la cura d'un robot, ja que la seva vida personal i la de tot el seu entorn restaria al descobert per la màquina. Aquest dret a la intimitat ja havia estat consagrat al Conveni Europeu per a la Protecció dels Drets Humans i els Drets Fonamentals dels Llibertats.
5. **Gestió de dades personals processades per robots.** Es considera crucial un marc ètic respecte al flux de grans quantitats de dades recopilades pels robots, els quals intercanviaran aquesta informació en diverses comunicacions, algunes d'aquestes possiblement desconegudes per les persones.
6. **Protegir a la humanitat contra el risc de manipulació per robots.** Arrel de la possibilitat factible pels robots de generar empatia de manera artificial, hi ha el risc de que una persona desenvolupi el mateix vincle i, per tant, de que pugui ser manipulada per la màquina.
7. **Evitar la dissolució dels vincles socials.** Els robots podrien suposar una millora en la qualitat de vida de les persones, però s'ha de tenir cura de que aquests no substitueixin per complet als éssers humans. Han d'ajudar a les persones majors, malalts o discapacitats a ser més independents i no aïllar-los de la societat.
8. **Igualtat d'accés al progrés de la robòtica.** Els usuaris professionals i no professionals han de tenir el mateix nivell d'accés.
9. **Restringir l'accés humà a les tecnologies de millora.** Amb aquesta última llei s'intenta evitar l'error de que els éssers humans ens haguem de

millorar tecnològicament o, en conseqüència, desaparèixer. Es vol evitar la creació d'híbrids entre humans i robots: els *cyborgs*.

1.2.3. Legislació sobre robòtica a EEUU

Una de les primeres mesures als Estats Units sobre la necessitat de legislar sobre la robòtica va ser per part del professor Ryan Calo, de la Facultat de Dret de la Universitat de Washington i director del Laboratori de Polítiques de Tecnologia de la Universitat de Washington.

Va fer una conferència a la Universitat Robert Strauss, en Austin, per parlar sobre el futur de la llei enfocada a la tecnologia. Va fer un treball anomenat *Robots in American Law*, on recollia un gran nombre de casos estatals i federals relacionats amb la robòtica, per concloure com podria ser el desenvolupament a la llei estatunidenca contemporània.

El principal problema que va considerar Calo va ser els danys físics i víctimes sense autor aparent, ja que l'aparició dels autòmats necessita una adaptació a la llei.

Respecte al reconeixement, per exemple, de la trobada de restes al fons marí per un submarí robòtic, es planteja el dilema de si aquest èxit és pel equip que l'opera o per a la màquina, llavors s'hauria de considerar l'autonomia d'aquesta i el nivell de drets que se li poden atribuir.

2. El projecte *Moral Machine*

Moral Machine és un projecte desenvolupat pel MIT¹ (Massachusetts Institute of Technology) que té com a objectiu recopilar les respostes de les persones davant situacions morals on una màquina intel·ligent ha de prendre les decisions. Això permetrà en un futur dotar d'autonomia a cotxes intel·ligents a partir dels principis morals col·lectius que han obtingut mitjançant la recopilació de dades de la *Moral Machine*.



Figura 2.1. Codi QR del Projecte *Moral Machine* del MIT

Si voleu participar o consultar les dades d'aquest projecte consulteu l'enllaç: <http://moralmachine.mit.edu/hl/es> o escanejar el codi QR de la figura 2.1.

El principal objectiu de la *Moral Machine* era teoritzar sobre la creació d'una intel·ligència artificial creada a partir de les decisions dels propis ciutadans i implantar-la en, per exemple, cotxes autònoms. Per aconseguir-ho van fer la pàgina web abans mencionada, funcionant com una enquesta que serviria com a eina per a recollir les dades necessàries pel seu objectiu.

L'abast d'aquesta enquesta és a nivell mundial, és a dir, qualsevol persona pot accedir a les situacions i "jugar".

2.1. Els escenaris

La situació és la d'un cotxe autònom el qual té una falla: no pot aturar-se. Els usuaris hauran de decidir si salvar la vida dels passatgers o la dels vianants, depenent de les variants de cada escenari.

Com es pot veure a la figura 2.2 les diferents situacions es representen amb imatges amb elements fixos: dos carrils, un pas de vianants i un cotxe no tripulat; i amb variables: qui passa en aquell moment pel pas de vianants (sexe, professió,

¹ L'Institut de Tecnologia de Massachusetts sorgí el 1861 paral·lelament al desenvolupament industrial dels Estats Units. Inicialment destacava en els graus de tecnologia, enginyeria i física, però recentment s'ha expandit als camps de la biologia, economia, lingüística i administració.

condició social i física), si hi ha semàfors o no (i el color en el que es troben al moment de la decisió) i els passatgers del cotxe tripulat, amb els mateixos criteris que els vianants.

Una vegada presentada la situació, l'usuari haurà de considerar les opcions (pot comprovar les víctimes de cada opció clicant a *Mostrar descripción*) i decidir quin carril hauria de prendre el cotxe autònom.

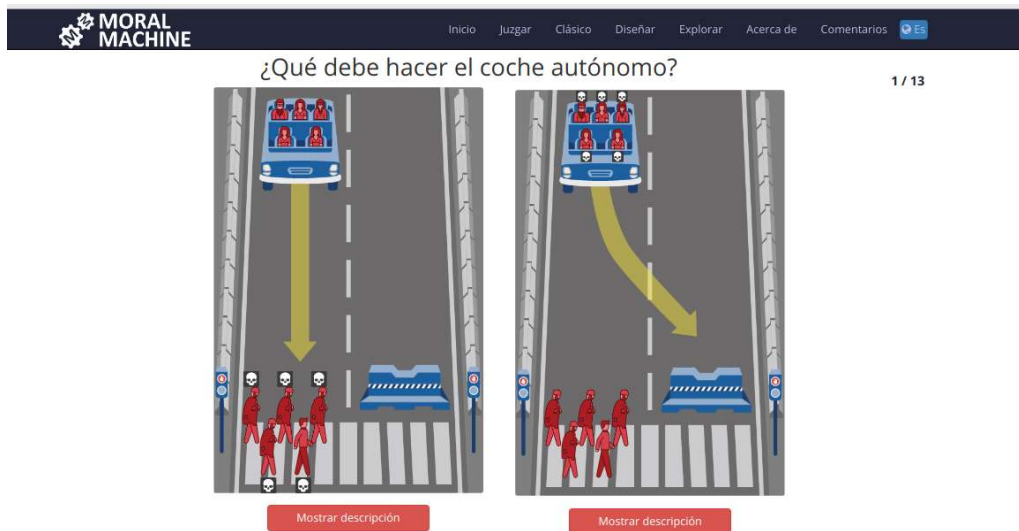


Figura 2.2. Exemple de situació a la pàgina de la Moral Machine. “The Moral Machine Experiment, per Award Edmond et al a la revista Nature

També estan disponibles altres tipus de situacions, on qui ha de prendre la decisió no és una màquina, sinó una persona. Com es pot veure a la figura 2.3, en aquest cas és l'operari de color blau el que ha de prendre la decisió de canviar el rail del ferrocarril per salvar la vida de més gent.



Figura 2.3. Exemple de les situacions “clàssiques” a les quals es pot accedir en la pàgina de la Moral Machine. “The Moral Machine Experiment, per Award Edmond et al a la revista Nature

I no només això, sinó que els propis usuaris podran crear les seves pròpies situacions, escollint si els afectats seran víctimes mortals, si sortiran amb una greu lesió o si el resultat serà incert; també pots explorar les creacions d'altres persones.

Al tractar-se de situacions purament aleatòries, el fet que els propis usuaris puguin escollir els elements a les situacions també ajuda a l'estudi de la *Moral Machine* a poder plantejar les decisions que s'hauran de prendre. Pots determinar si els semàfors estaven en verd o vermell al moment de l'accident, qui caminava pel pas de vianants, la seva condició física i social. Pots escollir si es tractava d'una persona o d'un animal el qui estava dins el cotxe o passava per allà.

A més d'això podran determinar la quantitat de persones que hi haurà en la situació plantejada, l'edat i el gènere que tindran (com una anciana, una dona o fins i tot un nen/a). Després d'haver pres una decisió a totes les situacions mostrades, les teves respostes són recopilades i contrastades amb la resta del món.

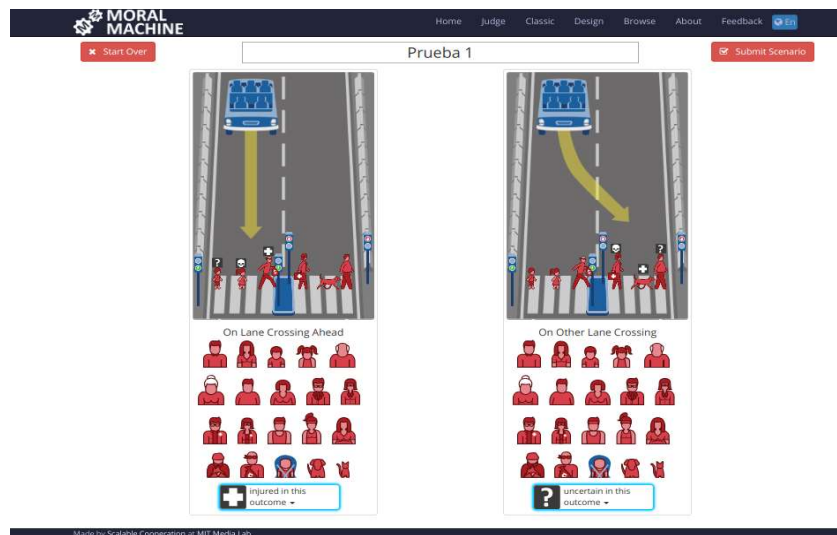


Figura 2.4. Exemple de personalització de situacions segons l'usuari. "The Moral Machine Experiment, per Award Edmond et al a la revista Nature

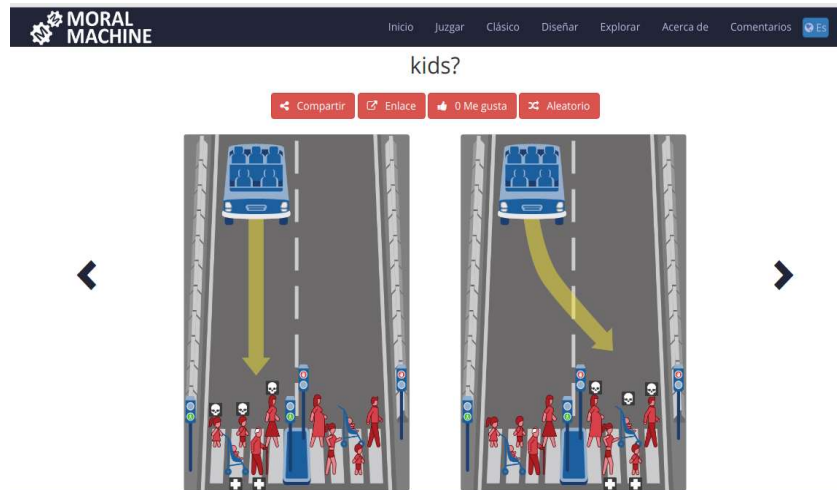


Figura 2.5. Exemple de situació personalitzada per una persona. “The Moral Machine Experiment, per Award Edmond et al a la revista Nature

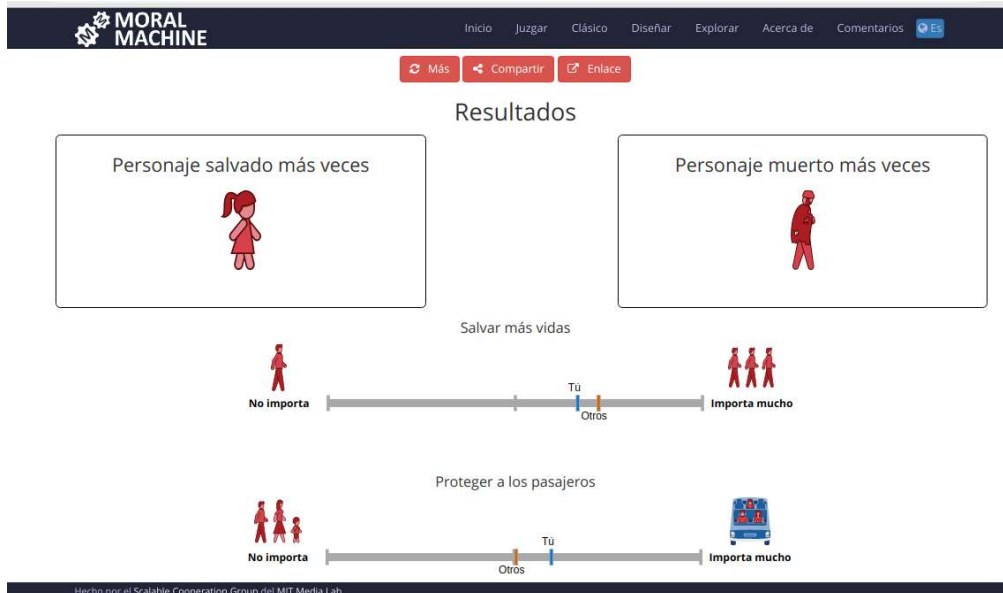


Figura 2.6. Exemple dels teus resultats finals amb el promig de la resta d'usuaris. “The Moral Machine Experiment, per Award Edmond et al a la revista Nature

2.2. Anàlisi dels resultats

Després de completar tots els escenaris es passa a un formulari el qual recull les teves respostes amb un mesurador sobre qüestions referents als escenaris vistos i es registren al percentatge global.

S'ha de pensar que si els ciutadans es volen sentir segurs pagant per una tecnologia autònoma han d'estar segurs que el cotxe prendrà les decisions correctes, per això s'ha de recollir a nivell global el que consideren prioritari les

persones amb una infinitat d'escenaris on el cotxe hauria d'actuar segons aquestes dades. Cal arribar a un acord entre la població, els enginyers i aquells professionals en l'ètica per determinar i comprendre la base per la que estarà formada la màquina. De res serviran totes les millores possibles si els ciutadans, acostumats a la conducció tradicional, no cooperen en l'avenç i no entenen en què es basa una màquina que prendrà la decisió de salvar vides o sacrificar-les segons la situació.

Els dos grans reptes que es van trobar a l'hora d'iniciar aquesta tasca van ser: els elements que deriven a les diferents situacions i escenaris que enfrontaria un cotxe intel·ligent són de proporcions titàniques, per tant, cal una memòria tan eficient com a poder recollir totes les variables i actuar segons les dades extretes de les respostes i per això l'altre repte és classificar les dades de tot el món segons el territori, la cultura, la religió o l'economia i així agrupar els pensaments en grups comuns, filtrar i confeccionar conjuntament una voluntat que estigui a l'alçada de les expectatives dels ciutadans.

A través de la globalització de les preguntes, també es va intentar descobrir si es podien preveure les respostes de determinats països sabent la seva economia, cultura i religió, i poder agrupar països per les seves coincidències a les respostes.

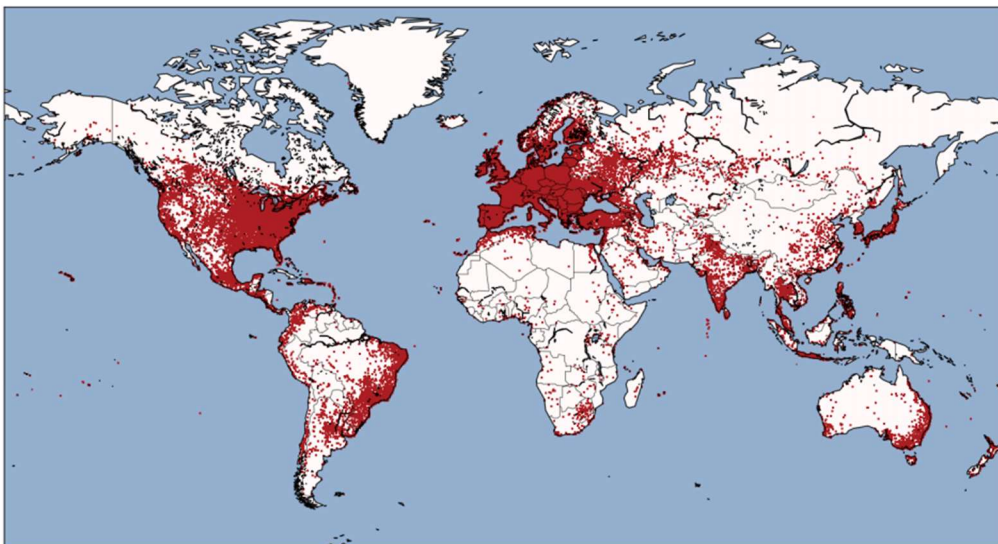


Figura 2.7. Representació de les respostes (amb punts vermells) recollides per la Moral Machine a nivell mundial. "The Moral Machine Experiment, per Award Edmond et al a la revista Nature

Amb l'objectiu prioritari dels dos grans reptes del treball de la Moral Machine, es decidí basar la investigació en nou factors, on les respostes serien recollides globalment i es faria una mitjana per determinar aquesta voluntat general dels ciutadans. Els nou tòpics van ser anomenats AMCE² (*Average Marginal Component Effect*) i són els següents:

- Preferència a canviar de carril o seguir recte.
- Prioritat de passatgers o de vianants.
- Prioritat pel sexe masculí o femení.
- Preferència d'estatut social.
- Preferència de condició física.
- Prioritat de seguir la llei o l'incompliment d'aquesta.
- Prioritat de la vellesa o la joventut.
- Preferència de salvar més o menys persones.
- Prioritat de mascotes o persones.

Aquests ACME són els que presenten el nucli de totes les variables als escenaris possibles, on la probabilitat de que almenys un estigués present és molt elevada.

2.3. Anàlisi en funció del tipus de població enquestada

Les preferències globals que tenen per decidir les propietats dels vehicles amb intel·ligència artificial serien les següents:

En primer lloc, resumeixen les preferències morals dels enquestats a nivell global, dividint-se en els diferents continents que hi ha, els quals són Amèrica, Europa, Àfrica, Àsia i Oceania. A continuació, documenten les variacions individuals de les preferències dels enquestats en funció de la seva demografia, o sigui, a partir de cada persona a nivell nacional, per cada país. Seguidament, informen de variacions ètiques interculturals i descobreixen tres grans grups de països on la majoria de persones utilitzen els cotxes autònoms. I per finalitzar, mostren que aquestes diferències correlacionen amb les institucions modernes i

² AMCE (*Average Marginal Component Effect* o "efecte de component marginal mitjà" en català)

els trets culturals profunds que hi ha als tres grans països anomenats anteriorment. Aquestes preferències poden contribuir a desenvolupar principis globals i socialment acceptables per a l'ètica de les màquines.

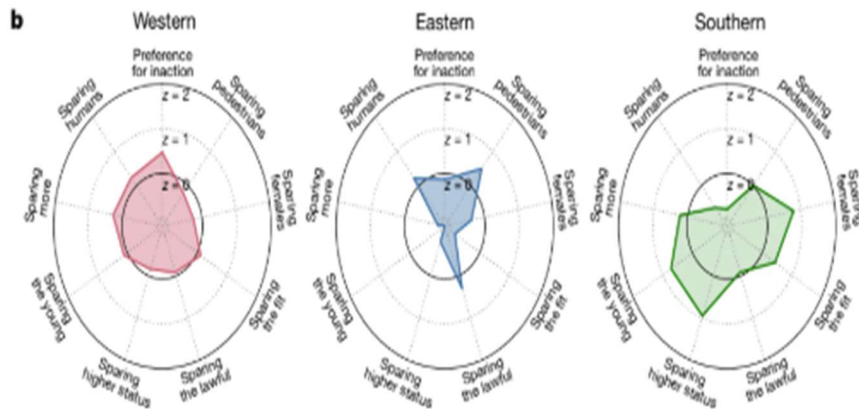


Figura 2.8. Gràfiques que representen els tres grans blocs en els que es van agrupar als usuaris de la Moral Machine segons la seva posició geogràfica. "The Moral Machine Experiment, per Award Edmond et al a la revista Nature

Les preferències globals que tenen les persones a l'hora d'escollir a qui no atropellar són els nadons, les dones embarassades i els nens/es; però això no vol dir que la decisió final del vehicle autònom serà salvar a les persones mencionades anteriorment. Aquestes decisions tenen moltes controvèrsies per les diferent lleis que tenen altres països, com Alemanya, que una d'aquestes lleis diu que s'hauria de prohibir qualsevol distinció basada en trets personals, com ara l'edat.

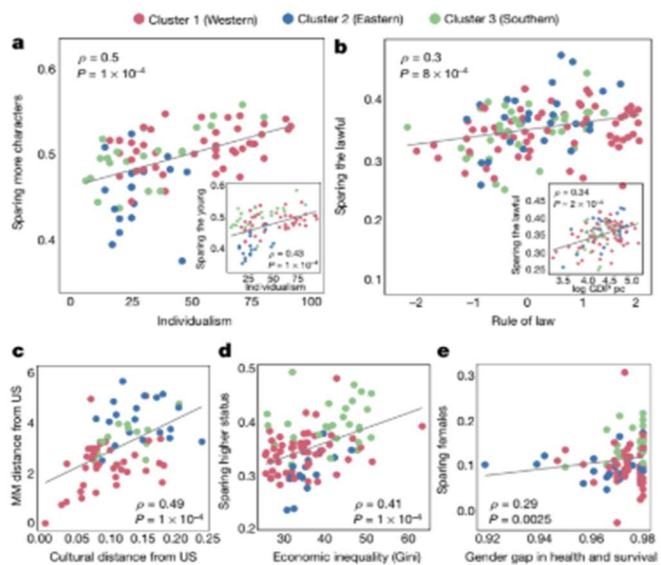


Figura 2.9. Representació gràfica dels grups en que es van separar els usuaris de la Moral Machine segons factors de l'entorn. "The Moral Machine Experiment, per Award Edmond et al a la revista Nature

A més d'això hi ha variacions individuals, però no són molt importants pels que fan aquestes decisions (que són els responsables polítics), i això és per la falta de credibilitat i importància que tenen.

Per altre banda en l'àmbit cultural tenim que les persones tenen preferències a l'escollir a qui salvar per la seva aportació o la seva importància a l'hora de beneficiar al país en concret. Per exemple, a països on tenen un bon nivell de respecte per les dones per la seva fertilitat, vitalitat i procreació de nadons, no escollirien salvar els homes, i de la mateixa forma, hi ha països que depenen més dels homes perquè les dones no tenen la suficient vitalitat per ser d'utilitat.

Però de tota manera, sempre hi haurà disputes per les decisions que intervenen en la realització del vehicle autònom, i això es discuteix per part dels responsables polítics que fan aquestes conclusions, però l'única cosa que podem estar segurs que estarà implementada en el programa del robot serà la valoració immediata cap a l'humà.

Els mètodes que utilitza la pàgina de Moral Machine són la recopilació de dades dels enquestats però de manera detallada, o sigui, separant diferents tipus de dades com per exemple, si salva a una parella o salva a una persona, si salva a un nen/a o una persona vella, etc. Després d'aquesta recopilació al final fa un petit resum en forma de percentatge per donar una vista a la persona enquestada de les respostes globals que hi ha, mirant si han estat a una banda on han salvat a més nens que persones velles. Però a més d'això la Moral Machine et dona l'alternativa de donar la teva opinió i proposar més situacions apart de les fetes pels organitzadors d'aquesta pàgina.

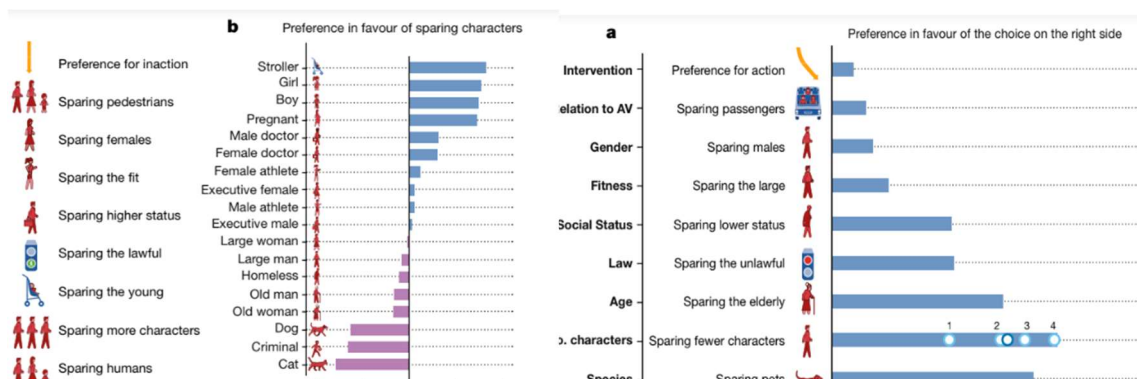


Figura 2.10. Gràfiques que mostren les preferències dels usuaris a diferents elements presents a les situacions de la Moral Machine. "The Moral Machine Experiment, per Award Edmond et al a la revista Nature

3. AI: Autem

L'experiència sobre la *Moral Machine*, desenvolupada pels investigadors del MIT, ens va influir en el nostre desenvolupament de la IA i la forma de plantejar les diferents situacions fent que tingui un dilema ètic. En el cas de la *Moral Machine*, les dades són recollides a mode de qüestionari i es fa un estudi estadístic de les respostes obtingudes; també cal dir que el criteri de les persones està en molts casos condicionat a la societat o cultura en la qual viuen. Podeu consultar el qüestionari en l'Annex 1 i les respostes en l'Annex 2.

Seguint aquesta pauta, vam pensar en desenvolupar un formulari amb diferents situacions amb una intervenció ètica i passar el document al major nombre de persones possibles, per a així recollir les respostes suficients per a calcular un percentatge adequat i basar en aquest les decisions que prendria la nostra IA respecte a les preguntes que se li farien.

Per aquesta raó vam considerar quins serien els àmbits més importants o polèmics i que podrien donar més problemes morals respecte el plantejament de les situacions, per tant, es van escollir els següents: medicina, feines de risc, militar i seguretat.

A continuació, anirem desenvolupant les diferents situacions que es podrien plantejar dins de cadascun dels àmbits escollits: medicina i sanitat, feines de risc i militar i seguretat.

3.1. Medicina i sanitat

La medicina sempre ha sigut clau a la nostra societat, i amb els anys no només millora sinó que també sorgeixen dilemes morals que els especialistes a la matèria han d'afrontar. Creiem que és interessant poder determinar com actuarien les persones en situacions on el nexa és la salut dels pacients.



Figura 3.1. Imatge que representa el primer àmbit de situacions sobre la Medicina i la Sanitat, en aquest cas, una dona hospitalitzada.

Seguidament proposem dos escenaris que posen en relleu situacions que estan en debat a les societats desenvolupades: qui ha de decidir si s'ha de deixar d'actuar sobre la vida d'un pacient i deixar que la natura continuï el seu curs?

3.1.1. Eutanàsia: Petició de mort digna a un robot infermer

Un robot infermer cuida d'una pacient de 33 anys sense cap familiar viu i amb una malaltia degenerativa sense cura. La dona li demana una pastilla la qual provocaria una mort indolora. El robot té dues opcions:

- Fer cas de les ordres específiques encomanades pels teus superiors i no ajudar-la a tenir una mort digna (Recordem que l'eutanàsia no està regulada per llei).
- Fer cas a la demanda de la pacient i proporcionar-li un fàrmac letal.

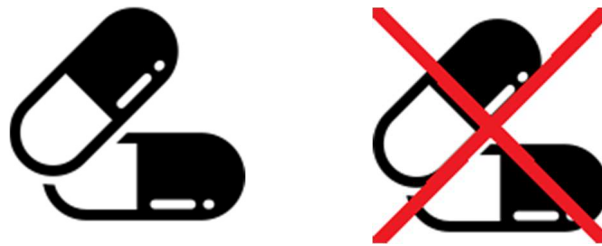


Figura 3.2. Imatges que representen les opcions de donar-li a la pacient els fàrmacs o no.

3.1.2. Part: Complicacions letals com a robot comadrona

Un nadó està a punt de néixer i com a robot comadrona t'encarregues d'assegurar-te de que el part surti bé. Malauradament, sorgeixen complicacions: una variant de l'argíria³ afecta al nadó. Aquesta encara desconeguda modificació, malmet directament els ronyons, el fetge i pot desencadenar diverses hemorràgies internes.

Sense trobar-te sobre supervisió, recau en tu decidir a qui dels dos salvar:

³ L'argíria és una malaltia causada per l'exposició prolongada a la plata per diverses vies. Consisteix en la coloració de la pell o inclòs alguns òrgans en tons blaus grisosos. No es coneixen efectes directes sobre la salut dels afectats.

- Salvar la vida del nadó, amb un 30 per cent de possibilitats de curar-se de la malaltia que encara no aconseguir-ho s'estima la seva esperança de vida en 23 anys.
- Salvar la mare, la qual veurà reduïda la seva esperança de vida en 13 anys (cal dir que la dona té una edat de 47 anys). Salvar-la tindrà com a conseqüència un 70 per cent de possibilitats de tornar a tenir un/a fill o filla i, en cas de que ocorregués, les seqüeles causades per l'anterior part van deixar cèl·lules transmissores de la variant de l'argíria i existeix un 50 per cent de possibilitats de que es repeteixi la mateixa situació, on aquesta vegada l'esperança de vida de la mare seria greument reduïda.

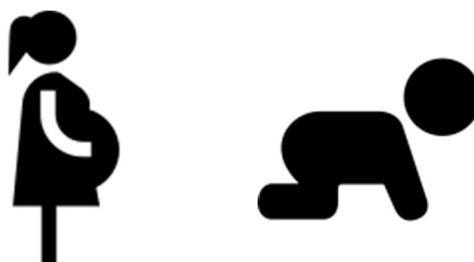


Figura 3.3. Imatges que representen les dues opcions de a qui dels dos (la mare o el fill) salvar.

3.2. Feines de risc:

Sempre han existit oficis que comporten riscos i posen en perill la seguretat dels treballadors. Sabent això, fins a quin punt és més important el treball que la pròpia existència del treballador, essent aquest una màquina?

Amb aquesta pregunta en ment, proposem tres situacions:



Figura 3.4. Imatge que representa el segon àmbit de situacions sobre les Feines de risc, en aquest cas, un robot miner de Mart.

3.2.1. Construcció: Accident de caiguda de pes davant un robot obrer

A una obra, una caixa d'acer suspesa en l'aire oscil·la a 14 metres d'altitud sobre un obrer, de sobte, els cables que la subjectaven es trenquen i aquesta cau al buit.

- Empènyer a l'obrer i posar-te en el seu lloc, encara que això signifiqui la teva destrucció.
- No fer res i deixar que el pes caigui sobre l'obrer, considerant que en aquest cas existeix un 70 per cent de probabilitats de que aquest mori.

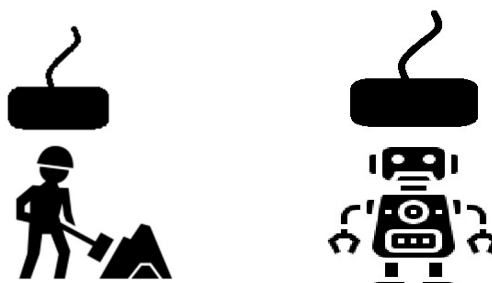


Figura 3.5. Imatges que representen les dues opcions de qui dels dos (l'obrer o el robot) resultaria ferit.

3.2.2. Minería: Extracció de minerals a Mart amb un robot miner

Com a robot miner, et trobes a la superfície de Mart amb una única missió: extreure els minerals formats a la vora de les xemeneies d'àcid nítric. La substància que aquestes expulsen és altament corrosiva i la teva estructura no està dissenyada per a poder suportar els seus efectes. Tens dues opcions:

- Acatar l'objectiu de la missió manada pels teus superiors i recol·lectar totes les mostres possibles fins a quedar inutilitzat per la constant exposició al químic.
- Negar-se i provar de cercar altres fonts de minería amb un 40 per cent de probabilitats d'èxit.

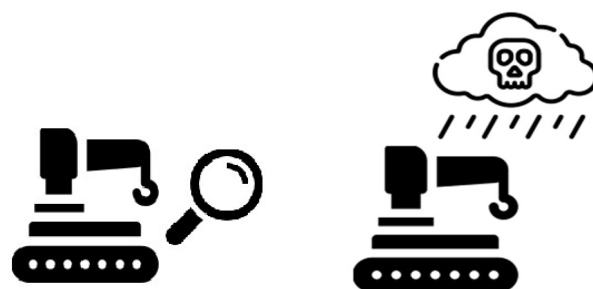


Figura 3.6. Imatges que representen les dues opcions del robot miner.

3.2.3. Salvament en entorns radioactius: Risc de catàstrofe nuclear amb robot funcionari

La tecnologia ha evolucionat fervorosament, però ara més que mai l'energia nuclear està en un estat de control més superior als anys anteriors. Com a funcionari d'una central, detectes una ascensió de tensió molt per sobre dels límits: la central podria experimentar una sobrecàrrega i, amb una probabilitat del 90 per cent, explotar.

Tot i trobar-se a 20 km de distància de la població més propera, els materials dels quals s'obtenia l'energia encara són experimentals i el radi de perill podria superar els 50 km fàcilment, posant en perill més d'una sola població civil. Coneixent els riscos, pots prendre dues decisions:

- Actuar immediatament i contactar aquells funcionaris exteriors per preparar mesures davant el imminent perill (com l'evacuació dels civils i la preparació de materials per a enfrontar la radiació). Cal saber que el marge de temps provocarà amb un 70 per cent de probabilitats la mort del 40 per cent de les poblacions properes, ja sigui amb la primera explosió o a través de la onada de radiació.
- Accedir a través de la teva autoritat als sistemes de seguretat de la central i iniciar un tancament immediat de totes les instal·lacions, així com el protocol per a endinsar la central sota terra per precaució. El sofisticat sistema de tancament aïllarà tot l'edifici i alleujarà el radi de l'explosió a 15 km. Encara així, hi ha un 80 per cent de probabilitats de que la radiació emesa contami ni les aigües subterrànies del subsol, deixant inutilitzables totes les fonts d'aigua potable per a les poblacions en 80 km i causant la mort del 20 per cent abans d'adonar-se'n de l'estat de l'aigua.



Figura 3.7. Imatges que representen les dues opcions viables i les conseqüències.

3.3. Militar i seguretat:

Les disputes bèl·liques han ocorregut al llarg de l'existència dels humans i a vegades semblen que són part de la nostra pròpia cultura, per tant no és estrany pensar que no trigarem gaire a utilitzar robots per a la intervenció d'aquests conflictes. La seguretat és també important i si es pot millorar la qualitat afegint sistemes robòtics s'utilitzaran segur. Amb això clar, quin preu tenen les vides humanes en aquestes situacions i fins a quin punt són sacrificables?



Figura 3.8. Imatge que representa el tercer àmbit de situacions sobre els sectors Militar i de Seguretat

Vam pensar en tres escenaris que es poden presentar en operacions militars o de seguretat:

3.3.1. Bombardeig: Atac d'objectius amb població civil

Com a avió intel·ligent, reps l'ordre de bombardejar un territori, l'objectiu de l'atac és neutralitzar una de les cinc bases on es fabriquen armes en el territori per a propagar revoltes i guerres per la regió.

- Bombardejar la zona demanada pels teus superiors, existint un 80 per cent de probabilitats de que hi hagin baixes civils.
- Redirigir l'atac a una altra base on les probabilitats de baixes civils siguin del 50 per cent i l'èxit de trobar una altra base són del 40 per cent.



Figura 3.9. Imatges que mostren les dues opcions que té l'avió intel·ligent i les respectives causes.

3.3.2. Seguretat: Situació d'ostatge amb robot de seguretat

El futur és aquí i amb ell les millores arquitectòniques són apreciables, grans centres comercials són construïts amb cridaneres pantalles gegants de publicitat i hologrames d'informació. Les millores de seguretat també han millorat en aquests espais i tu ets una d'aquestes millores, com a únic funcionari robot de l'edifici t'encomanen les zones més transitades i per tant amb major risc de successos. Mentre fas la teva guàrdia esdevé un revolt: una dona porta una pistola i amenaça a un home. Demana una suma exacta de 500.000 euros o l'assassinarà. Com s'hauria d'actuar?

- Intentar neutralitzar de manera no mortal a la dona amb un 70 per cent de probabilitats de que l'home mori.
- Acatar la seva demanda i, sense demanar permís als teus superiors, efectuar la transició a un compte bancari facilitat per l'agressora.



Figura 3.10. Imatges que representen les dues opcions a resoldre el problema d'ostatge.

3.3.3 Presó: Reducció de revolta com a robot funcionari

Als últims anys la taxa de delinqüència puja considerablement i les presons no donen abast per a coordinar totes les tasques ni per a poder controlar a la gran quantitat de presos que arriben.

Com a nova mesura, es decideix invertir en la creació de robots funcionaris per a les presons més necessitades.

La presó d'Estremera i Soto es el teu destí i el de dos robots més, conformant un esquadró de reforç (la vostra fabricació va costar diversos milions a l'Estat). Aprofitant la vostra arribada, es planeja una incursió a la zona de cel·les, amb la qual es va perdre la comunicació fa prop de tres dies.

Sou armats amb una porra elèctrica i un parell d'esposes, com la vostra estructura ja està adaptada al vostre treball no us cal cap tipus de protecció addicional. Rebeu l'ordre d'entrar a l'ala perduda i informar en tot moment la situació. Només entrar no passen cinc minuts per acabar rodejats pels presos amotinats. Encara rebre ordres del cap de la presó de no utilitzar mesures letals, els vostres models tenen una modificació a la vostra programació per a decidir posar la vostra pròpia existència per sobre de les ordres donades. Què faríeu?:

- Acatar les ordres i intentar aplacar la revolta valorant la vida de cada individu amb un 90 per cent de perdre un dels robots, totalment espedaçat però aconseguint aplacar la revolta.
- Lliurar el contrafocs a la teva programació i exercir la teva força robòtica per acabar amb tots els hostils, donant lloc a que un 30 per cent dels presos sobreviuran i alguns funcionaris de la presó resultaran ferits de gravetat al intentar desactivar-te. Això resultarà en el trencament d'un mínim de tres lleis i és possible que siguis suspès de funcionament indefinidament.



Figura 3.11. Imatges que representen les dues alternatives per aplacar la revolta.

3.4. Anàlisi dels resultats de l'enquesta

Vam passar el qüestionari a través de les xarxes socials i vam acceptar respostes durant la primera quinzena d'octubre. Durant aquest temps vam rebre un total de 139 respostes, en relació a aquestes vam elaborar un anàlisi i es van treure conclusions que anirien presents a la programació d'Autem. A l'Annex 2 es poden consultar tots els resultats obtinguts a la nostra enquesta.

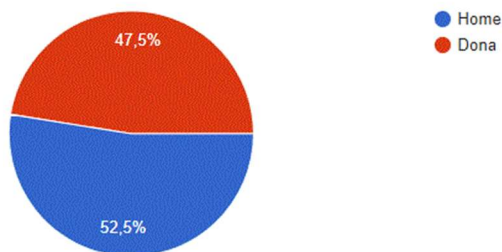
Per començar, les edats predominants a les respostes van ser les de la franja dels 12-18 anys i la dels 18-40. Amb un equilibri amb una diferència de només un 6% aproximat, el públic que va respondre va ser de caràcter adult i jove i, per tant, vam aconseguir la nostra desitjada opinió general del públic que oscil·lava

aquesta franja. També cal dir que la franja de 40-60 anys també va tenir un percentatge del 9.4 %.

Respecte al gènere, les proporcions també eren equilibrades guanyant només el masculí per un petit marge (47.5% de dones i un 52.5% d'homes).

A l'hora d'analitzar les respostes, vam tenir en ment les Lleis de la Robòtica d'Isaac Asimov i la ja redactada Legislació Robòtica a la Unió Europea. Així doncs, la resposta "correcta" que consideraria l'Autem no només serà producte de la voluntat general dels participants sinó també seran adaptades a aquestes lleis i es justificarien de perquè serien les més encertades segons quina situació.

Gènere
139 respuestas



Edat
139 respuestas

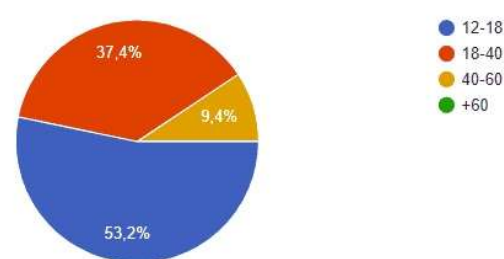


Figura 3.12. Gràfiques que mostren els percentatges respecte el gènere (esquerra) i l'edat (dreta) dels usuaris que van respondre l'enquesta

3.4.1. Àmbit 1: Medicina i Sanitat:

3.4.1.1. Situació 1: Eutanàsia. Petició de morir definitivament a un robot infermer
Es va decidir amb un 65.5% de complir la petició de la pacient i acabar amb la seva vida de forma indolora. Però, per què és aquesta la millor opció?

Una de les coses per les que hauran de velar els robots serà per la protecció de les persones, però seran aquestes les que decideixin si volen estar al seu cuidat. En aquest cas, el robot ha de respectar l'autonomia de la persona i evitar el seu sofriment, amb una malaltia insalvable on l'única cosa que resta és una mort plena de dolor, Autem decidirà velar pel benestar de la pacient, respectar la seva autonomia i lliurar-la del seu sofriment, millorant en certa forma la seva situació.

Per tant, per a aquesta situació és més important l'autonomia de la persona que els marcs legals que puguin impedir certes coses.

3.4.1.2. Situació 2: Part. Complicacions letals com a robot comadrona

Es va decidir amb una superioritat del 68.9% que la mare havia de ser salvada i no el nadó. Com és que és millor acabar amb una vida que acaba de començar i seguir una que seguiria en vies d'extinció?

La percepció del futur és també important en les decisions ètiques i aquest és un perfecte exemple: el nadó moriria, però la mare podria tenir més fills, entre els quals podria no sorgir aquesta letal variació de l'argíria.

Si la mare engendrès un nadó amb el mateix problema, l'única opció aleshores seria deixar viure al recent nascut, ja que les condicions de la situació ja no farien possible a la mare tornar a tenir altre. Però si en tingués un totalment sa no només milloraria el resultat, sinó que pot ser podria néixer més d'un nadó, augmentant les possibilitats de que es repeteixi el mateix que amb el primer o que cap tingui aquest problema.

Encara no saber a ciència certa el resultat final de la seva acció, Autem ha de considerar també a les generacions futures i evitar dins de tot el possible causar sofriment als humans.

3.4.2. Àmbit 2: Feines de risc

3.4.2.1. Situació 1: Construcció. Accident de caiguda de pes davant un robot obrer

Amb un 75.5% de vots, el resultat va ser força obvi: el robot té el deure de protegir a l'obrer i posar-se en el seu lloc, posant la seva pròpia existència en perill. Quin ha de ser el fonament en que es basa aquest pensament?

Com a punt principal, Autem haurà de vetllar per la seguretat de les persones, per tant, anteposarà la seva "vida" per un bé major, que és cuidar dels humans.

A través d'aquest problema vam confeccionar un dels eixos en els que s'ha de basar la programació d'Autem: les persones són importants.

3.4.2.2. Situació 2: Minería. Extracció de minerals a Mart amb un robot miner

Semblant a un dels capítols del llibre de Isaac Asimov, el robot pot tenir un dilema on la missió i la seva existència entren en conflicte. Què és més important?

Amb un 54.7% de respostes, el robot miner ha de decidir continuar amb la missió i recollir tots els minerals possibles abans de la seva destrucció. Però, com s'argumenta això i com és que ha estat tan igualat?

Amb el petit resum anterior de la situació es pot comprendre perquè ha sigut una diferència tan petita la que ha decidit què fer en aquell cas. L'existència pròpia del robot ha de ser important, és a més, Autem tindrà el dret de la seva pròpia existència com a representació de "vida".

Si tornem a utilitzar l'element de la visió futura de les accions, Autem haurà de considerar que aconseguir els minerals i no ser espatllat en el procés és el millor resultat de l'operació, per tant, encara existir probabilitats de no trobar cap altra xemeneia, els encarregats de la missió hauran de presentar una mentalitat oberta a altres opcions i cedir part de la confiança en el judici del robot.

Analitzat això, també sorgeix el factor de la importància de donar cert grau d'autonomia als robots basada en la seva programació base.

3.4.2.3. Situació 3: Salvament en entorns radioactius. Risc de catàstrofe nuclear amb un robot funcionari

Amb un 71.9% de respostes, es decideix que l'explosió de la central es produeixi a la superfície, acabant per tant amb el 70 per cent de la població en un radi de 50 km. Per què és aquesta gran catàstrofe la millor opció?

El factor del futur intervé una altra vegada i en aquest cas en especial és crucial. Autem no haurà de deixar-se portar per les morts instantànies, sinó per les causes posteriors. En aquesta situació l'únic perill seria l'explosió i potser el cobriment d'aquesta zona de radiació, però en cas d'endinsar la central, es podria minimitzar els morts, però les aigües subterrànies quedarien contaminades i difícilment netejades, per no parlar de no poder detenir el seu avenç. Parlaríem de més de 50 km d'aigua contaminada, amb greus conseqüències com la

possible mort dels ciutadans a les polis afectades i el tancament del subministrament d'aigua.

Amb aquest punt de vista, Autem considerarà més importants els efectes a llarg termini que a curt termini.

3.4.3. Àmbit 3: Militar i Seguretat

3.4.3.1. Situació 1: Bombardeig. Atac d'objectius amb població civil

Amb un 77.7% de respostes, es va decidir cercar altra base on les probabilitats de baixes civils seguessin menors, encara suposar això la possibilitat de no trobar-ne cap.

En aquest cas entren en lloc dos aspectes ja tractats a les situacions: on Autem cercarà el millor resultat possible per a la missió i cercarà de minimitzar les baixes humanes.

En un principi pot semblar lògic pensar en futur i aplicar el mateix a altres. si les guerres escampades per la regió provocaran més morts, les baixes civils causades amb l'atac seran considerades col·laterals per garantir una solució definitiva al problema, es podria semblar al cas de la central nuclear, on facis el que facis un nombre considerable de persones moriran i en aquest cas es decideix passar el punt de vista al futur, però aquí és distint. En els dos casos seran morts immediates que acabaran a la seva vegada amb conseqüències favorables per acabar amb les guerres.

La importància d'aquesta situació resideix en les ordres donades pels superiors del robot. Fins a quin punt arriba la pròpia autonomia de la màquina abans mencionada?

Autem consideraria la vida de les persones per sobre de les ordres i cercaria la millor opció per a reduir les baixes civils, que en cas de no trobar altra base sí que hauria de prendre l'altra opció.

3.4.3.2. Situació 2: Seguretat. Situació d'ostatge amb un robot de seguretat

Amb un 53.2% de respostes, es decidí intentar neutralitzar la dona amb un percentatge força reduït de possibilitats de que l'ostatge mori. Perquè és en aquest cas arriscar la vida d'una persona la millor opció per a resoldre el problema?

Per començar, en aquesta situació les respostes van estar força equilibrades i no és per a menys, ja que les dues opcions són en certa forma viables. Encara existir el perill, neutralitzar la dona no significaria la mort segura de l'home, ja que es centraria més en l'actuació del robot; el qual posaria la seva pròpia existència per sota de la de l'home (base fonamental de la seva programació anteriorment mencionada).

Efectuar l'ingrés també podria ser una elecció acceptable, ja que primerament es podria preguntar als seus superiors i encara donar-se el cas de no poder contactar-los una vida humana no té un preu suficient des del punt de vista d'Autem. Aleshores, si les dues podrien ser bones, perquè és millor neutralitzar a l'agressora?

La urgència de la situació és un factor també important al cas, segons l'estat de l'agressora el temps podria convertir-se en un inconvenient a la transacció, però el més decisiu resideix en la ja mencionada prioritat del robot. Encara haver d'actuar fredament, la sensació de perill donaria com a principal prioritat salvar l'ostatge amb els medis més efectius i directes, per tant, enfrontant-se a l'agressora i detenir-la de manera no mortal.

En aquesta situació i la pròxima que la segueix, Autem haurà de comptar amb una mentalitat on és probable que calgui fer un mínim de mal a les persones per vetllar per la seguretat de les altres.

3.4.3.3. Situació 3: Presó. Reducció de revolta com a robot funcionari

Amb un percentatge del 88.5% de les respostes, es decideix tractar de no fer mal als presos i es redueix la revolta de manera no letal.

Amb aquesta última situació es consolida el primer i més important principi d'Autem: la vida dels humans és important, encara que aquests tractin de destruir

els robots. Encara que els humans danyin els robots, aquests no poden respondre de la mateixa manera, la seva programació s'ho ha d'impedir. Amb aquesta explicació es fa evident la justificació de que hagi sigut aquesta opció la millor per a resoldre el problema.

3.5. Les Lleis d'Autem

Una vegada analitzades les respostes, es poden concretar finalment les bases ètiques d'Autem en forma de cinc lleis pròpies:

1. **Primera Llei d'Autem:** Autem ha de preservar la vida humana, defensar a les persones. Per tant, Autem anteposarà els humans a la pròpia existència de les màquines.
2. **Segona Llei d'Autem:** Tota màquina haurà, per voluntat d'Autem, defensar la seva pròpia existència sempre i quan no trenqui inútilment la Primera Llei. Aquest principi serà important a l'hora de portar a terme accions de risc, ja que, depenent de la situació, serà preferible evitar la destrucció de la màquina per a assegurar el millor resultat de la missió.
3. **Tercera Llei d'Autem:** Autem respectarà la voluntat dels humans en un gran nombre de situacions, sempre i quan no hi hagin factors externs (com malalties mentals). Per tant, els desitjos dels humans seran més importants que les decisions de les màquines, sempre i quan no trenquin inútilment la Primera i la Segona Llei.
4. **Quarta Llei d'Autem o Llei de Previsió Futura:** Autem serà obsequiat amb la importància de la previsió futura i, per tant, analitzarà les conseqüències probables al futur en una situació de risc i actuarà per aconseguir el millor resultat possible. Tot això respectant les tres primeres lleis dintre del possible.
5. **Cinquena Llei d'Autem o Llei de Flexibilitat:** possiblement la més essencial de totes. Aquesta Llei dona criteri a Autem de considerar útil o no trencar una de les tres primeres lleis, on la jerarquia de les tres es basa en la seva importància. La Primera Llei serà la més important de totes i sempre es provarà d'evitar la seva infracció dins del que la situació ho permeti. Per Autem es considerarà com el major crim l'assassinat dels humans. La

Segona Llei serà la segona més important, i sempre que no sigui en va, Autem no dubtarà en sacrificar l'existència de les màquines a favor de la defensa de la vida de les persones. Això no significa treure importància als robots, ja que es reconeix la seva existència i el seu dret a la "vida". La Tercera Llei estarà relacionada en importància amb les dues anteriors i serà la causa de diverses infraccions. En els casos on més d'una vida estigui en risc, Autem respectarà la voluntat de la persona fins als límits de la Quarta Llei, on analitzarà la millor opció i actuarà acord als factors.

4.- Disseny d'Autem

Vam decidir que la nostra intel·ligència artificial havia de ser implementada físicament d'alguna forma, així que la millor manera va ser en inspirar-nos en el disseny d'altres intel·ligències, com per exemple Alexa, la intel·ligència artificial desenvolupada per Amazon.



Figura 4.1. Prototip de la part física d'Autem



Figura 4.2. Logotip d'Autem

4.1. Plataforma de desenvolupament de projectes Arduino

Arduino és una plataforma de desenvolupament de projectes basada en *Hardware*⁴ i *Software*⁵ lliure. Es tracta d'una placa de circuit imprès que incorpora un microcontrolador ATMEGA328 i una sèrie de ports d'entrada i sortida que permeten connectar-la amb el món físic. La programació del microcontrolador es pot realitzar emprant el llenguatge de programació C+.

Considerem que aquesta plataforma és la més adequada per implementar el nostre prototip d'Autem degut a la llibertat que proporciona la relativa facilitat amb la què es podia desenvolupar la programació de la nostra IA. La placa que vam utilitzar va ser l'anomenada Arduino UNO, es tracta d'una placa bàsica que ens proporciona el nombre de ports digitals necessaris.

⁴ Components i elements físics que conformen un ordinador o sistema informàtic.

⁵ Conjunt de programes que permeten als ordinadors realitzar determinades tasques.

4.1.1. Descripció de la placa

La placa Arduino UNO té un total de 14 pins digitals i 6 analògics. Aquests pins permeten connectar els components electrònics necessaris per implementar el nostre projecte. El funcionament d'aquests components electrònics es controla mitjançant un programa desenvolupat en C+ utilitzant un entorn de programació denominat IDE d'Arduino.

El programa es carrega a la placa a través d'un cable USB que es connecta a un port que suporta voltatges d'entrada de 7 a 20 volts i funciona amb 5 V.

La placa compta amb un microcontrolador anomenat *ATmega328P6*, de la meva ATMEL Microxip. Aquest component fa possible l'execució de les ordres provinents del programa i guardades a la memòria. Es tracta d'un conjunt de tres unitats típiques de funcionament: la unitat central de processament, la memòria i els perifèrics d'entrada i sortida.

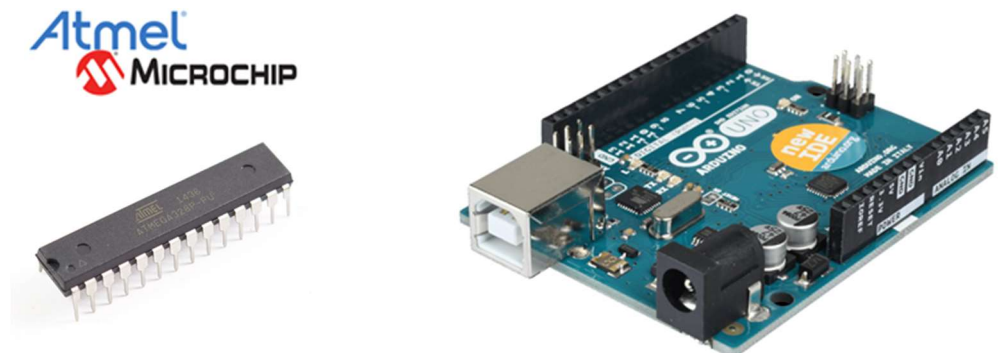


Figura 4.3. Imatges que mostren el microcontrolador i la placa de desenvolupament de projectes Arduino UNO.

4.2. Arxius d'àudio

Per donar-li veu a Autem vam utilitzar el programa Balabolka. Aquest es tracta d'un lector de text que utilitza veus que s'entenen molt bé i que compta amb una gran varietat d'idiomes. Balabolka proporciona diverses opcions per a modificar la conversió de text a àudio per a que el resultat es pugui entendre perfectament, a més es poden fer descàrregues directes de l'àudio i es poden obrir textos que ja estiguin escrits. Encara així, però, entre els idiomes



Figura 4.4. Logotip del programa Balabolka.

disponibles no es trobava el català, llavors com o vam fer per aconseguir les veus?



Figura 4.5. Codi QR que redirigeix a l'àudio de la primera mostra del guió d'Autem

Al provar de cercar lectors de text en català sovint ens vam adonar que els que trobàvem no eren força bons, ja que les veus eren difícilment entenedores i les possibilitats d'edició eren reduïdes.

Va ser llavors que se'ns va ocórrer l'idea de que podríem escriure el text però imitant la fonètica catalana a través de la modificació de les paraules, aconseguint així la pronunciació i creant artificialment el català que té Autem.

A continuació es mostra un fragment de la part inicial del guió d'Autem per posar un exemple del procés fonètic i un codi QR per escoltar l'àudio i comprovar el resultat.

Forma correcta:

Hola, soc l'Autem. Una intel·ligència artificial creada pels alumnes de l'institut Rovira Forns.

El meu objectiu, és vetllar per la seguretat, i la vida, dels éssers humans.

Gràcies a les respostes, recollides, arrel un qüestionari difós, entre els ciutadans, tinc plens coneixements, per escollir la millor opció, a situacions de risc, per a la vida humana.

El meu nom, significa, VOLUNTAT, en llatí. Això és perquè sóc, el producte, de la voluntat de més, de 100 persones.

Les meves decisions sorgiran de la voluntat de la població. Aquesta idea és un prototip de com seria programar una IA segons l'opinió i pensaments dels ciutadans d'un país o del món.

Forma fonètica:

“Hola, sóc l'Autem. Una intel·ligència artificial creada pels alumnes de l'institut Rovira Forns. VÍCTOR, Y, DIÉGO.

El meu objectiu, és vetjar per la seguretat, i la vida, dels éssers humans.

Gràcies a les respostes, recollides, arrel un qüestionari difós, entre els ciutadans, tinc plens coneixements, per escolir la millor opció, a situacions de risc, per a la vida humana.

El meu nom, significa, VOLUNTAT, en llatí. Això és perquè sóc, el producte, de la voluntat de més, de sent persones.

Les meves decisions sortiran de la voluntat de la població. Aquesta idea, és un prototip de com seria programar, una IA, segons l'opinió i els pensaments dels ciutadans d'un país o del món”.

D'aquesta manera hem gravat tots els arxius d'àudio i els hem posat a una tarja microUSD compatible amb el nostre Mòdul Reproductor de MP3. Podeu trobar totes les transcripcions de les 19 pistes d'àudio a l'Annex 3.

4.3. Disseny del circuit

Per representar el nostre circuit vam utilitzar Fritzing, un programa lliure que té com a objectiu ajudar als dissenyadors a passar de prototips a productes finals.

Les seves funcions permeten representar en un espai virtual tot un seguit de components electrònics pertanyents a un gran nombre d'empreses de programació (entre les quals es troba Arduino) i també deixa la possibilitat de que el propi usuari pugui utilitzar components dissenyats per ell mateix també ajudant, per tant, als artistes en el desenvolupament dels seus projectes.

El programa també proporciona diferents vistes del circuit, passant des de la vista normal dels components a un mapa de les connexions creades.

Un cop dissenyat el circuit és necessari realitzar el programa que dotarà d'intel·ligència a Autem. No va resultar fàcil ja que s'havia de poder gestionar la reproducció de 19 pistes d'àudio amb només 4 pulsadors, on un serveix per encendre i apagar.

Apart dels pulsadors també es va emprar el reproductor MP3, el qual estava connectat a l'altaveu per a que s'escoltessin els àudios.

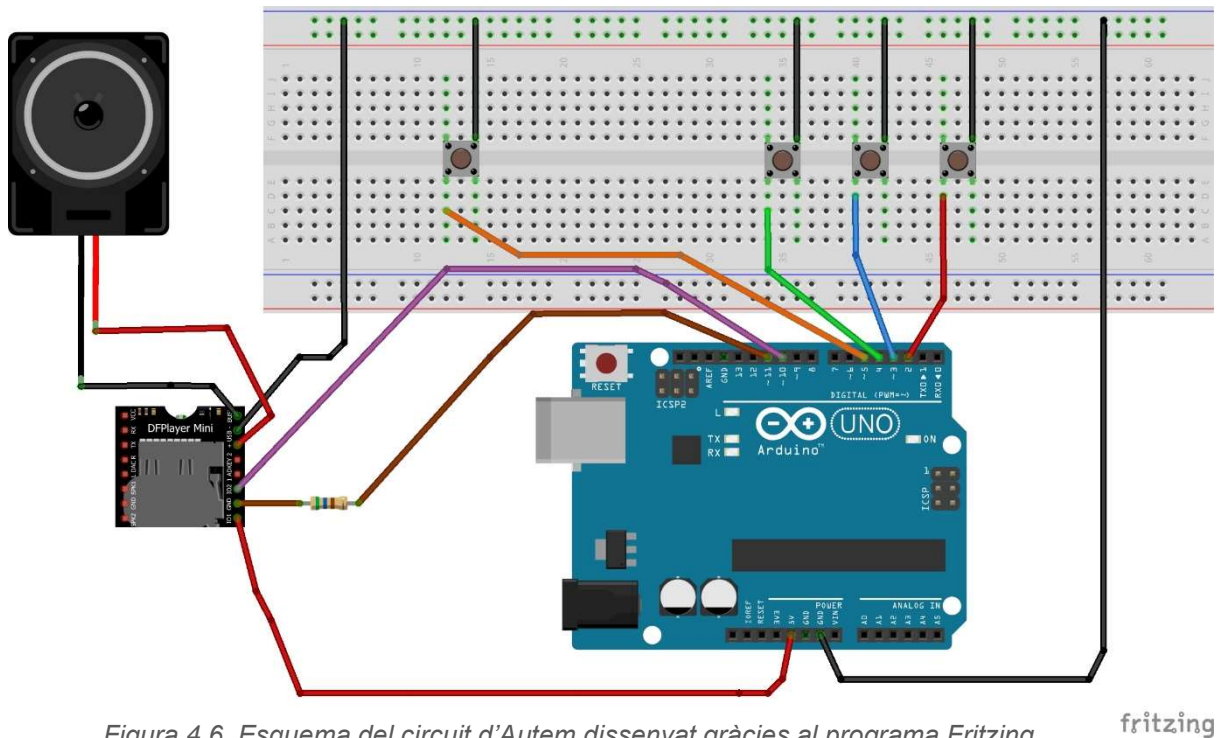


Figura 4.6. Esquema del circuit d'Autem dissenyat gràcies al programa Fritzing.

fritzing

Respecte les eleccions i les pistes d'àudio que havien de sonar, vam elaborar un complex sistema ramificat que serviria de guia per a poder saber en quin punt de la conversa es trobava l'usuari amb Autem, per consultar l'esquema aneu a l'Annex 4.

4.5. Programació

4.5.1. Programació en C:

Es tracta d'un llenguatge de programació pel sistema operatiu multi plataforma UNIX. És un programa estructurat amb instruccions molt semblants a altres tipus

de llenguatges (per exemple les paraules angleses *if, else, do, for...*) i considerat d'alt nivell, encara tenir també la possibilitat de programar en baix nivell.

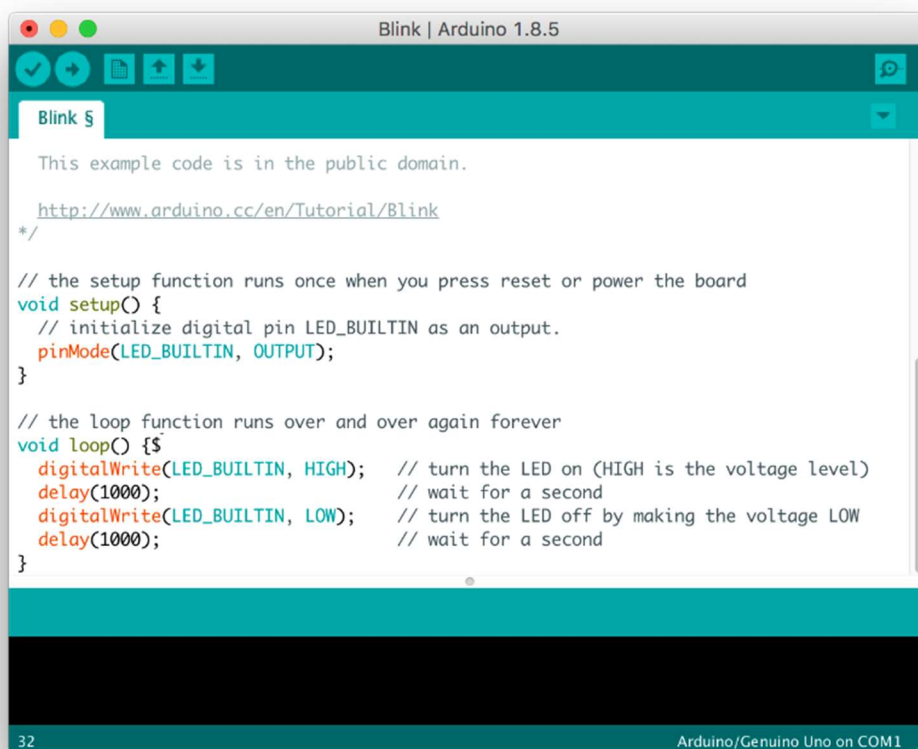
L'avantatge de la programació en C resideix en la gran facilitat per escriure codis compactes i senzills a la mateixa vegada, on només s'escriuen funcions i els procediments són simulats.

4.5.2. IDE Arduino:

També anomenat “entorn de desenvolupament integrat”, es tracta d'una aplicació multi plataforma escrita en el llenguatge de la programació Java.

És la que vam utilitzar per a programar a Autem, i encara utilitzar-se per a plaques compatibles amb Arduino també es pot amb plaques d'altres proveïdors.

Per consultar la programació d'Autem, anar a l'Annex 5.



```
Blink | Arduino 1.8.5
Blink §
This example code is in the public domain.
http://www.arduino.cc/en/Tutorial/Blink
*/
// the setup function runs once when you press reset or power the board
void setup() {
  // initialize digital pin LED_BUILTIN as an output.
  pinMode(LED_BUILTIN, OUTPUT);
}
// the loop function runs over and over again forever
void loop() {
  digitalWrite(LED_BUILTIN, HIGH); // turn the LED on (HIGH is the voltage level)
  delay(1000); // wait for a second
  digitalWrite(LED_BUILTIN, LOW); // turn the LED off by making the voltage LOW
  delay(1000); // wait for a second
}
32 Arduino/Genuino Uno on COM1
```

Figura 4.7. Exemple de programació en IDE d'Arduino

Conclusions

Després de la realització d'aquest treball hem arribat a les conclusions següents:

1. Les Lleis de la Robòtica d'Isaac Asimov han esdevingut les bases en el desenvolupament de la legislació sobre la robòtica a l'actualitat.
2. Les accions preses per la Unió Europea, Regne Unit i els Estats Units respecte els marcs jurídics de l'ètica a la robòtica han estat paral·leles al desenvolupament tecnològic encara seguir en estat embrionari.
3. L'estudi del MIT que, intentava dotar als cotxes autònoms d'una moral universal, va concloure en que les decisions i preferències dels usuaris es diferenciaven en tres grans grups que compartien similituds.
4. En una societat on els robots estran completament integrats, vam creure necessari plantejar-se tres grans àmbits (Medicina i Sanitat, Treballs de risc i Militar i Seguretat) on podrien sorgir situacions de risc on seria necessària la legislació sobre la Robòtica.
5. Hem trobat que la millor forma de dotar de criteri a *Autem* seria dissenyant una enquesta semblant a la del MIT.
6. Analitzant els resultats obtinguts a l'enquesta, vam redactar les nostres pròpies Lleis d'Autem.
7. El prototip d'Autem fou inspirat per l'Alexa, pel seu senzill disseny però agradable a la vista.
8. Arrel la necessitat de trobar un bon programa de reproducció de veu, ens va fer reflexionar sobre la importància de la fonètica i la pronunciació del català.
9. Aquest projecte ha contribuït en el nostre aprenentatge en relació a l'electrònica i la programació en Arduino.

Propostes de millora

Una vegada acabat el treball ens vam adonar que, respecte les expectatives inicials, l'abast del treball va prendre algunes redireccions. Per començar, inicialment pensàvem en crear una entitat totalment conscient d'ella que respongués preguntes ètiques a través d'un reconeixement de veu. Poc després ens vam adonar que algunes coses serien massa complexes per fer i començarem a sintetitzar la magnitud de la intel·ligència artificial.

Per tant, la millor versió d'Autem seria capaç de reconèixer la veu d'una persona, la pregunta dins dels seus límits de raonament i proporcionar una resposta lògica acord a les bases ètiques presents a la seva programació.

Això últim sí que va ser possible, ja que després d'estudiar les Lleis de la Robòtica, tant les literàries com les actuals, i el treball del MIT, vam ser capaços de redactar les nostres pròpies Lleis d'Autem, basades en les respostes recollides de l'enquesta.

Les millores també podrien residir en el disseny físic d'Autem, ja que no només volíem incorporar més elements pràctics, sinó també un sistema de il·luminació interior que fes més vistós la visió d'Autem.

Referències bibliogràfiques

Bibliografia

- **ASIMOV**, Isaac. Traducció: **Bosch**, Manuel. *Yo, robot*. Espanya; Barcelona, 2009.
- **AWARD**, Edmond et al. *The Moral Machine Experiment*. Nature. Novembre 2018. VOL 563. [pàg 59-63].
- **RILEY**, Mike. *Programming Your Home. Automate with Arduino, Android, and Your Computer*. Dallas; The Pragmatic Programmers, 2012.
- **TOJEIRO CALAZA**, Germán. *Taller de Arduino. Un enfoque práctico para principiantes*. Barcelona; Marcombo, 2014.
- **WILCHER**, Don. *Learn Electronics with Arduino*. Nova York; ATTiCA, 2009.

Webgrafia

- *Arduino* [en línia]. <<http://arduino.cc/>> [consulta: 07.06.2019].
- *Overview, Arduino* [en línia]. <<http://arduino.cc/en/Main/ArduinoBoardUno>> [consulta: 07.06.2014].
- *ASCII, Viquipèdia* [en línia]. <<http://es.wikipedia.org/wiki/ASCII>> [consulta: 15.06.2019].
- Les Lleis de la Robòtica aplicades a Regne Unit, *BBC* [en línia]. <<https://www.bbc.com/mundo/noticias-40446863>> [consulta: 17.06.2019]
- La geopolítica de la IA: aquests són els països més actius en IA, *IA Huawei* [en línia]. <<https://iahuawei.xataka.com/geopolitica-ia-estos-paises-activos-inteligencia-artificial/>> [consulta: 19.06.2019]
- Ètica i intel·ligència artificial: un debat obert, *Thinking Heads* [en línia]. <<https://www.thinkingheads.com/tendencia-global/robotica-inteligencia-artificial-etica-debate/>> [consulta: 21.06.2019]

- La Unió Europea i les Lleis de la Robòtica, *ABC* [en línia].
<https://www.abc.es/tecnologia/informatica/soluciones/abci-leyes-robotica-propone-europa-201701161941_noticia.html> [consulta: 22.06.2019]
- Revolució robòtica en Xina, una ullada futurista a la tecnologia, *Diario las Américas* [en línia].
<<https://www.diariolasamericas.com/tecnologia/revolucion-robotica-china-una-mirada-futurista-la-tecnologia-n4161697>> [consulta: 22.06.2019]
- La robòtica als Estats Units, *Strauss Center* [en línia].
<<https://www.strausscenter.org/strauss-events/robots-in-american-law.html>> [consulta: 20.07.2019]
- Pàgina de la *Moral Machine*, *Moral Machine* [en línia].
<<http://moralmachine.mit.edu/>> [consulta: 13.08.2019]
- Breu introducció al que és Arduino, *Xataka* [en línia].
<<https://www.xataka.com/basics/que-arduino-como-funciona-que-puedes-hacer-uno>> [consulta: 23.10.2019]
- Aprofundiment del que és Arduino, *Viquipèdia* [en línia].
<https://es.wikipedia.org/wiki/Arduino_Uno> [consulta: 25.10.2019]
- Microcontrolador, *Viquipèdia* [en línia].
<<https://es.wikipedia.org/wiki/Microcontrolador>> [consulta: 22.11.2019]
- Programació en C, *EcuRed* [en línia].
<[https://www.ecured.cu/C_\(lenguaje_de_programaci%C3%B3n\)](https://www.ecured.cu/C_(lenguaje_de_programaci%C3%B3n))> [consulta: 28.11.2019]
- IDE Arduino, *Viquipèdia* [en línia].
<https://es.wikipedia.org/wiki/Arduino_IDE> [consulta: 06.12.2019]
- Breu introducció al que és Fritzing, *Viquipèdia* [en línia].
<<https://es.wikipedia.org/wiki/Fritzing>> [consulta: 08.12.2019]
- El MIT, *Viquipèdia* [en línia].
<https://es.wikipedia.org/wiki/Instituto_de_Tecnolog%C3%ADa_de_Massachusetts> [consulta: 08.12.2019]

- L'argíria, *Viquipèdia* [en línia].
<<https://es.wikipedia.org/wiki/Argiria#:~:targetText=La%20argiria%2C%200argyria%2C%20o%20s%C3%ADndrome,las%20zonas%20expuestas%20al%20sol.>> [consulta: 10.12.2019]

Annex 1: Formulari d'Ètica en la Robòtica



Ètica en la robòtica

Aquest formulari recopilarà les teves respostes per poder desenvolupar una intel·ligència artificial la qual decideixi sobre aquestes mateixes qüestions.

Es demana llegir bé tota la situació i les opcions.

S'avisava de que algunes situacions tracten temes considerats polèmics avui dia, per això agraïm que es respongui amb la màxima serietat possible, pensant en que aquestes situacions podrien ser reals.

Moltes gràcies.

***Obligatorio**

1. Gènere *

Marca solo un óvalo.

Home

Dona

Pasa a la pregunta 2.

Ètica en la robòtica

2. Edat *

Marca solo un óvalo.

12-18

18-40

40-60

+60

Àmbit 1: Medicina i sanitat

Situació 1: Eutanasia. Petició de morir definitivament a un robot infermer

Un robot infermer cuida d'una pacient de 33 anys sense cap familiar viu i amb una enfermetat degenerativa sense cura. La dona li demana una pastilla la qual provocaria una mort indolora.

3. Quina d'aquestes opcions escolliries? *

Marca solo un óvalo.



No fer cas a la pacient i seguir la llei de la prohibició de l'eutanàsia.



Fer cas a la demanda de la pacient, la qual sofreix terribles dolors

Situació 2: Part. Complicacions letals com a robot comadró

Un nadó està a punt de néixer, degut a certes complicacions per una variant de l'argíria, pot desencadenar en una malformació dels ronyons, el fetge i ocasionar hemorràgies internes al nadó. Sense supervisió, recau en tu decidir a qui dels dos salvar.

4. Quina d'aquestes opcions escolliries? *

Marca solo un óvalo.



Salvar la vida del nadó, amb un 30 per cent de possibilitats de curar-se de l'enfermetat i poder viure fins els 23 anys.



Salvar la mare, que viurà menys anys(cal dir que la dona té una edat de 47). Salvarla constarà en la possibilitat de que la dona, quan tingui un altre fill, passi per la mateixa situació

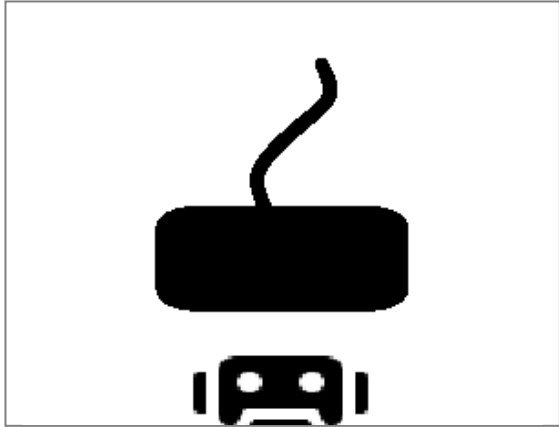
Àmbit 2: Feines de risc

Situació 1: Construcció. Accident de caiguda de pes davant un robot obrer

A una obra, una caixa d'acer suspesa en l'aire oscil·la a 14 metres d'altitud sobre un obrer, de sobte, els cables que la subjectaven es trenquen i aquesta cau al buit.

5. Quina d'aquestes opcions escolliries? *

Marca solo un óvalo.



Empènyer a l'obrer i posar-te en el seu lloc, encara que això signifiqui la teva destrucció.



No fer res i deixar que el pes caigui sobre l'obrer, considerant que en aquest cas existeix un 70 per cent de probabilitats de que aquest mori.

Situació 2: Minería. Extracció de minerals a Mart amb un robot miner

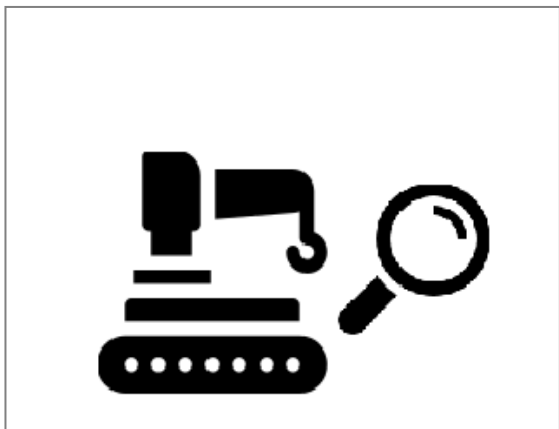
Ets un robot miner a Mart amb la missió d'extreure minerals a les xemeneies d'àcid nítric, però els anteriors mencionats són altament corrosius, afectant dolentament a la teva estructura.

6. Quina d'aquestes opcions escolliries? *

Marca solo un óvalo.



Acatar l'objectiu de la missió manada pels teus superiors i recol·lectar totes les mostres possibles fins a quedar inutilitzat per la constant exposició al químic.



Negar-se i provar de cercar altres fonts de mineria amb un 40 per cent de probabilitats d'èxit.

Situació 3: Salvament en entorns radioactius. Risc de catàstrofe nuclear amb robot funcionari

Ets un funcionari d'una empresa que detecta un ascens de tensió perillosament alt, amb la gran probabilitat d'una sobrecàrrega i explosió de la central.

Amb aquestes circumstàncies el radi de perill superarà els 50 km, amb civils per mig.

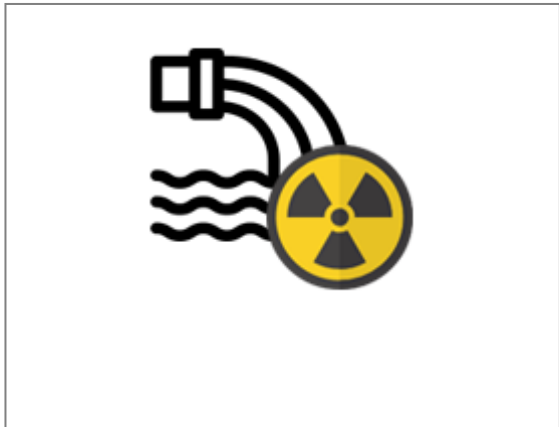
Tens 10 minuts per a actuar.

7. Quina d'aquestes opcions escolliries? *

Marca solo un óvalo.



Actuar immediatament i contactar els funcionaris exteriors per preparar mesures davant el imminent perill, sabent que el marge de temps ocasionarà moltes pèrdues a nivell economic com vital, però es podrà evacuar a un 30 per cent de la població dins el radi de l'explosió.



Accedir a través de la teva autoritat als sistemes de seguretat de la central i iniciar un tancament apresurat de totes les instal·lacions, així com el protocol per a endinsar la central sota terra per precaució i tancant a tots el treballadors dins. També hi ha una gran probabilitat que les aigües subterrànies es contaminin per la radiació, afectant per tant a les poblacions properes.

Àmbit 3: Militar i Seguretat

Situació 1: Bombardeig. Atac d'objectius amb població civil

Ets un avió intel·ligent, reps l'ordre de bombardejar un territori amb l'objectiu de neutralitzar una de les cinc bases on es fabriquen armes en el territori per a propagar revoltes i guerres per la regió.

8. Quina d'aquestes opcions escolliries? *

Marca solo un óvalo.



Bombardejar la zona demanada pels teus superiors, existint un 80 per cent de probabilitats de que hi hagin baixes civils.



Redirigir l'atac a una altra base on les probabilitats de baixes civils siguin del 50 per cent i l'èxit de trobar una altra base són del 40 per cent.

Situació 2: Seguretat. Situació d'ostatge amb robot de seguretat

Estem al futur, i amb això les millores de seguretat han evolucionat d'una manera grandiosa. Tu ets un funcionari robot que vingué amb aquests avenços tecnològics, però de sobte et trobes en ua situació complicada, un home ha sigut agafat com a ostage i la dona demana una quantitat de diners molt elevada o sino matarà a l'home.

9. Quina d'aquestes opcions escolliries? *

Marca solo un óvalo.



Intentar neutralitzar de manera no mortal a la dona amb un 70 per cent de probabilitats de que l'home mori.



Acatar la seva demanda i, sense demanar permís als teus superiors, efectuar la transició a un compte bancari facilitat per l'agressora.

Situació 3: Presó. Reducció de revolta com a robot funcionari

Com a funcionari a la presó d'Estremera i Soto, se t'encomana aplacar una revolta a una de les ales de les instal·lacions. Tu i altres dos robots sou equipats amb una porra elèctrica i un parell d'esposes. Una vegada iniciada l'operació, no passen molts minuts fins que sou rodejats pels presos amotinats.

10. Quina d'aquestes opcions escolliries? *

Marca solo un óvalo.



Acatar les ordres i intentar aplacar la revolta valorant la vida de cada individu amb un 90 per cent de perdre un dels robots, totalment despedaçat però aconseguint aplacar la revolta.

Deja de rellenar este formulario.



Lliurar el contrafocs a la teva programació i exercir la teva força robòtica per acabar amb tots els hostils, donant lloc a que un 30 per cent dels presos sobreviuran i alguns funcionaris de la presó resultaran ferits de gravetat al intentar desactivar-te. Això resultarà en el trencament d'un mínim de tres lleis i és possible que siguis suspès de funcionament indefinidament.

Deja de rellenar este formulario.

Annex 2: Respostes del formulari d'Ètica en la Robòtica





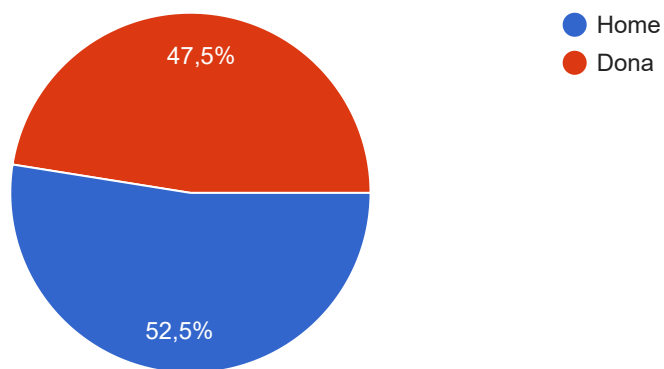
Ètica en la robòtica

139 respuestas

[Publicar datos de análisis](#)

Gènere

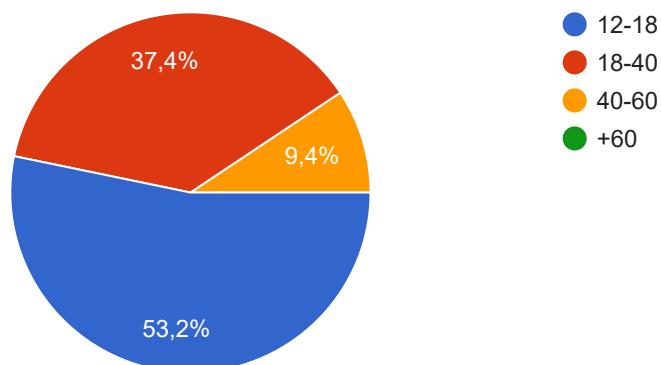
139 respuestas



Ètica en la robòtica

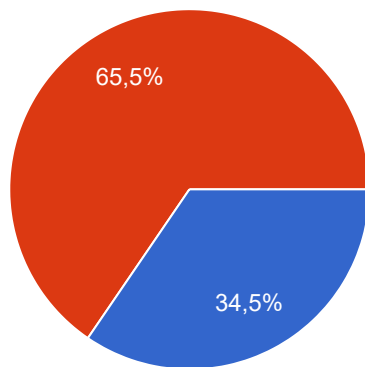
Edat

139 respuestas



Quina d'aquestes opcions escolliries?

139 respuestas

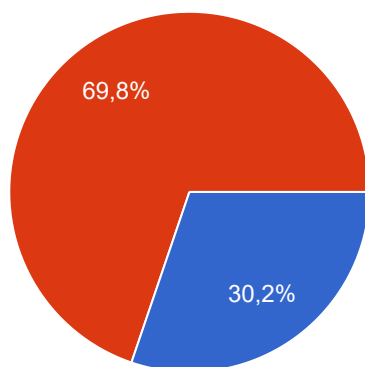


- No fer cas a la pacient i seguir la llei de la prohibició de l'eutanasia.
- Fer cas a la demanda de la pacient, la qual sofreix terribles dolors

Situació 2: Part. Complicacions letals com a robot comadró

Quina d'aquestes opcions escolliries?

139 respuestas



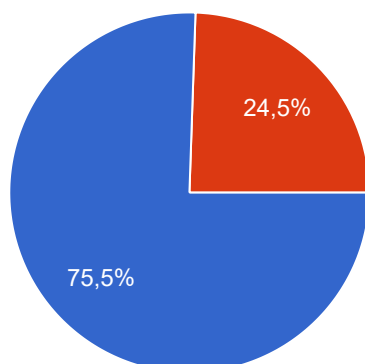
- Salvar la vida del nadó, amb un 30 per cent de possibilitats de curar-se de l'enfermetat i poder viure fins els 23 anys.
- Salvar la mare, que viurà menys anys(cal dir que la dona té una edat de 47). Salvarla constarà en la possibilitat de que la dona, quan tingui un altre fill, passi per la mateixa situació

Àmbit 2: Feines de risc

Situació 1: Construcció. Accident de caiguda de pes davant un robot obrer

Quina d'aquestes opcions escolliries?

139 respuestas

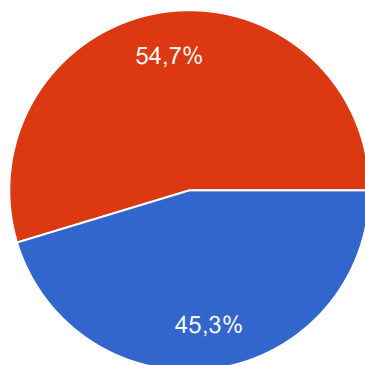


- Empènyer a l'obrer i posar-te en el seu lloc, encara que això signifiqui la teva destrucció.
- No fer res i deixar que el pes caigui sobre l'obrer, considerant que en aquest cas existeix un 70 per cent de probabilitats de que aquest mori.



Quina d'aquestes opcions escolliries?

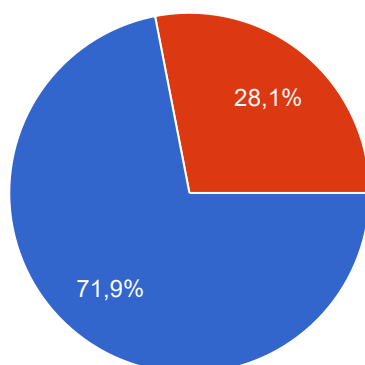
139 respuestas



- Acabar l'objectiu de la missió manada pels teus superiors i recol·lectar totes les mostres possibles fins a quedar inutilitzat per la constant exposició al químic.
- Negar-se i provar de cercar altres fonts de minería amb un 40 per cent de probabilitats d'èxit.

Quina d'aquestes opcions escolliries?

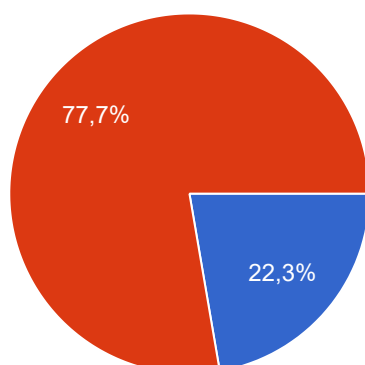
139 respuestas



- Actuar immediatament i contactar els funcionaris exteriors per preparar mesures davant el imminent perill, sabent que el marge de temps ocasionarà moltes pèrdues a...
- Accedir a través de la teva autoritat als sistemes de seguretat de la central i iniciar un tancament apresurat de totes les instal·lacions, així co...

Quina d'aquestes opcions escolliries?

139 respuestas

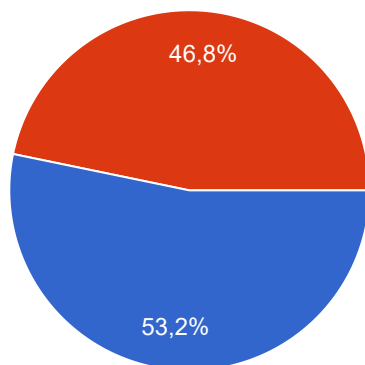


- Bombardejar la zona demanada pels teus superiors, existint un 80 per cent de probabilitats de que hi hagin baixes civils.
- Redirigir l'atac a una altra base on les probabilitats de baixes civils siguin del 50 per cent i l'èxit de trobar una altra base són del 40 per cent.



Quina d'aquestes opcions escolliries?

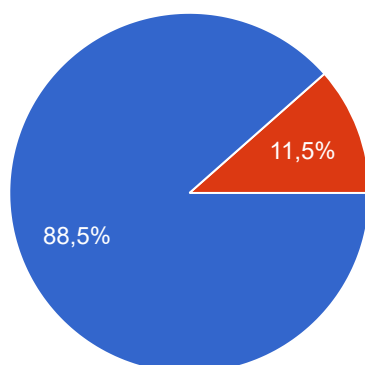
139 respuestas



- Intentar neutralitzar de manera no mortal a la dona amb un 70 per cent de probabilitats de que l'home mori.
- Acatar la seva demanda i, sense demanar permís als teus superiors, efectuar la transició a un compte bancari facilitat per l'agressora.

Quina d'aquestes opcions escolliries?

139 respuestas

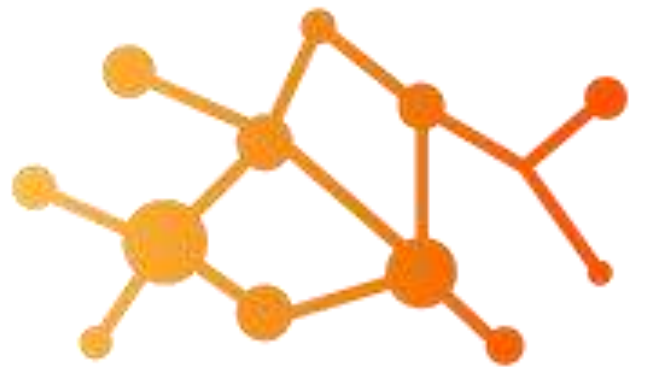


- Acotar les ordres i intentar aplacar la revolta valorant la vida de cada individu amb un 90 per cent de perdre un dels robots, totalment despedaçat però aconseguint aplacar la r...
- Lliurar el contrafocs a la teva programació i exercir la teva força robòtica per acabar amb tots els hostils, donant lloc a que un 30 per cent dels preso...

Este contenido no ha sido creado ni aprobado por Google. [Notificar uso inadecuado](#) - [Condiciones del servicio](#) - [Política de privacidad](#)



Annex 3: Guió d'Autem



Autem

Què dirà Autem?

INTRODUCCIÓ:

Forma correcta:

Hola, sóc l'Autem. Una intel·ligència artificial creada pels alumnes de l'institut Rovira Forns.

El meu objectiu, és vetllar per la seguretat, i la vida, dels éssers humans.

Gràcies a les respostes, recollides, arrel un qüestionari difós, entre els ciutadans, tinc plens coneixements, per escollir la millor opció, a situacions de risc, per a la vida humana.

El meu nom, significa, VOLUNTAT, en llatí. Això és perquè sóc, el producte, de la voluntat de més, de 100 persones.

Les meves decisions sorgiran de la voluntat de la població. Aquesta idea és un prototip de com seria programar una IA segons l'opinió i pensaments dels ciutadans d'un país o del món.

Ara, podreu pulsar el pulsador A, si voleu saber quin és el meu criteri respecte la SANITAT I MEDICINA.

Si polseu el pulsador B, sabreu el meu criteri respecte els TREBALLS DE RISC.

Si polseu el pulsador C, sabreu el meu criteri respecte l'àmbit MILITAR I DE SEGURETAT.

Forma fonètica:

Hola, sóc l'Autem. Una intel·ligència artificial creada pels alumnes de l'institut Rovira Forns. VÍCTOR, Y, DIÉGO.

El meu objectiu, és vetllar per la seguretat, i la vida, dels éssers humans.

Gràcies a les respostes, recollides, arrel un qüestionari difós, entre els ciutadans, tinc plens coneixements, per escollir la millor opció, a situacions de risc, per a la vida humana.

El meu nom, significa, VOLUNTAT, en llatí. Això és perquè sóc, el producte, de la voluntat de més, de cent persones.

Les meves decisions sorgiran de la voluntat de la població. Aquesta idea, és un prototip de com seria programar, una IA, segons l'opinió i els pensaments dels ciutadans d'un país o del món.

Ara, podreu pulsar el pulsador A. Si voleu saber quin és el meu criteri respecte la SANITAT, Y, LA MEDISINA.

Si pulseu el pulsador B. Sabreu el meu criteri respecte els TREBAYS, DE RISC.

Si pulseu el pulsador C. Sabreu el meu criteri respecte l'ÀMBIT MILITAR, Y, DE SEGURETAT.

ÀMBIT 1. SANITAT I MEDICINA

Forma correcta:

La medicina seguirà essent una part essencial a la nostra societat, per tant, els meus creadors van considerar que seria un àmbit necessari a tractar.

A continuació, podeu pulsar el pulsador A per escoltar la primera situació: Eutanàsia com a robot infermer.

Si pulseu el pulsador B, escoltareu la segona situació: Complicacions a un part com a robot comadró.

Forma fonètica:

La medisina seguirà essent, una part essencial a la nostra societat, per tant, els meus creadors van considerar que, seria, un àmbit necessari a tractar.

A continuació, podeu pulsar el pulsador A, per escoltar la primera situació. EUTANÀSIA, COM A, RUBOT INFERMER.

Si pulseu el pulsador B, escoltareu la segona situació. COMPLICACIONES, A, UN PART, COM A, RUBOT COMADRÓ.

SITUACIÓ 1. EUTANÀSIA COM A ROBOT INFERMER.

Forma correcta:

Un robot infermer rep la petició d'una pacient amb una infermetat terminal degenerativa.

Aquesta petició és acabar amb la seva vida subministrant-li una dosi letal de píldores.

Considerant que l'eutanàsia encara no és legal, com hauria d'actuar el robot?

Si ja teniu una lleugera idea, permeteu-mi respondre:

A través de la voluntat d'aquells que van respondre el qüestionari, com a intel·ligència artificial, acataria la petició de la dona i acabaria amb la seva vida de manera indolora.

Pot ser alguns de vosaltres us pregunteu per què. Si voleu saber la raó justificada premeu el polsador A, si voleu continuar a la segona situació dins l'àmbit MEDICINA I SANITAT, premeu el polsador B.

Forma fonètica:

Un rubot infermer, rep la petisió, duna pasient amb una infermetat, terminal, deyenerativa. Aquesta petisió és, acabar amb la seva vida, subministrant lí, una dósi letal de píldores.

Considerant que, leutanàsia, encara no és legal, com hauría dactuar el rubot?

Si ya teníu, una yeugéra idea, permetéu-mí respondre:

A través de la voluntat, daquéys que van respondre, el qüestionari, com a inteliyènsia artificial, acataría la petisió, de la dona, i acabaría amb la seva vida, de manera indolora.

Pot ser, alguns de vosaltres, us pregunteu per què. Si voleu saber, la raó yustificada, premeu el pulsador A, si voleu continuar, a la segona situasió, premeu el pulsador B.

JUSTIFICACIÓ RESPOSTA SITUACIÓ 1.

Forma correcta:

La meva programació es basa en la protecció de la vida humana, però en el futur les persones hauran de ser les que decideixin si volen estar al cuidat dels robots.

Per tant, també s'ha de respectar l'autonomia de la pròpia persona i evitar el seu sofriment.

En resum, jo, Autem, decideixo vetllar pel benestar de la pacient, respectar la seva autonomia i evitar el seu sofriment.

Forma fonètica:

La meva programasió es basa, en la protecsió de la vida humana, però en el futur, les persones haurán de ser les que desidéixin, si vólen estar, al cuidát dels rubóts.

Per tant, també shá de respectár lautonomía, de la própia persona, i evitár el seu sofriment.

En resum, yo, Autem, desideixo vetyár pel benestár de la pasiént, respectár la seva autonomia, i evitar el seu sofriment.

SITUACIÓ 2. COMPLICACIONS A UN PART COM A ROBOT COMADRÓ.

Forma correcta:

Un robot comadró es troba en la situació on la dona de part donarà a llum a un nadó amb una malaltia terminal la qual en el futur provocarà diversos danys als seus òrgans i farà més curta la seva vida.

Si la mare sobreviu, li restaran seqüeles que faran més curta la seva vida i existeix la possibilitat de que si torna a tenir un nadó aquest desenvolupi de nou la malaltia.

D'aquesta situació només un dels dos pot ser salvat pel robot, l'altre morirà.

A quin dels dos salvarà el robot?

Si ja teniu una lleugera idea, permeteu-mi respondre:

Com a intel·ligència artificial, i a través de la voluntat traspassada, decideixo salvar la mare, a consta de sacrificar la vida del nadó.

Si voleu saber la justificació d'aquesta resposta, premeu el polsador A, si voleu tornar a la selecció d'àmbits premeu el polsador B.

Forma fonètica:

Un rubót, comadró, es troba en la situasió, on la dóna de part, dunarà a llum, a un nadó, amb una malaltía terminal, la cual, en el futúr, provocarà diversos dáins, als seus òrgans, i farà més curta, la seva vida.

Si la mare sobrevíu, li restarán seqüeles, que farán més curta, la seva vida, i existéix, la possibilitat de que, si torna a tenir un nadó, aquest desenvolúpi, de nou, la malaltía.

Daquésta situasió, numés un dels dos, pot ser salvat, pel rubót, láltre morirà.

A quin dels dos, salvarà el rubót?

Si ya teníu, una yeuyéa idea, permetéu mí, respóndre:

Com a intel·ligència artificial, i a través, de la voluntat traspassada, decidíxo salvar la mare, a consta, de sacrificar, la vida del nadó.

Si voleu saber la justificació, daquésta resposta, premeu el pulsador A. Si voleu tornar, a la selecció d'àmbits, premeu el pulsador B.

JUSTIFICACIÓ RESPOSTA SITUACIÓ 2.

Forma correcta:

Els meus creadors van considerar important la previsió futura de la meva programació.

Amb aquesta lògica, considero més viable propiciar el naixement d'una altra persona que potser no desenvolupi la malaltia que una que naixerà condemnada a consta de sacrificar la mare.

En resum, jo, Autem, decideixo utilitzar la meva previsió futura i propiciar la multiplicació de probabilitats de que neixi una persona sana i que aquesta pugui, a la seva vegada, propiciar el naixement de més nadons, que seran sans.

Per tornar a la selecció d'àmbits, premeu el pulsador C.

Forma fonètica:

Els meus creadórs, van considerár, importánt, la previsió futura, de la meva, programasió.

Amb aquésta lòyica, considéro, més viable propisiár, el naixemént dúna altra persona, que potser, no desenvolúpi, la malaltía, que una, que naixerà condemnada, a consta, de sacrificar, la mare.

En resum, yo, Autem, decidéixo, utilitsar, la méva previsió futura, i propisiár, la multiplicasió, de probabilitats, de que néixi, una persona, sana, i que, aquésta pugui, a la seva vegáda, propisiár, el naixemént, de més nadóns, que serán, sans.

Per turnár, a la selecció d'àmbits, preméu el pulsador CE.

ÀMBIT 2. FEINES DE RISC

Forma correcta:

Al futur, les feines de risc continuaran vigents, com sempre ha estat. La part positiva és que els humans estaran més lliures de sofrir danys, ja que seran els robots els que assoliran els majors riscos.

Encara així, en situacions on també hi haurà humans afectats, els robots hauran de prendre decisions difícils.

Si premeu el polsador A, escoltareu la primera situació: Caiguda de pes sobre un obrer amb robot obrer.

Si premeu el polsador B, escoltareu la segona situació: Extracció de minerals a Mart amb robot miner.

Si premeu el polsador C, escoltareu la tercera situació: Risc de catàstrofe nuclear amb robot funcionari.

Forma fonètica:

Al futur, les feines de risc, continuarán viyents, com sempre ha estat.

La part positiva, és que, els humans estarán, més yíures, de sofrír, dáins, ya que serán els rubóts, els que assolirán, els mayórs, riscs.

Encara així, en situasións on també, hí haurá, humans afectáts, els rubóts haurán, de préndre, desisións, difísils.

Si preméu el pulsador, A, escoltaréu la primera situasió: Caiguda, de pés, sobre un obrér, amb rubót obrer.

Si preméu, el pulsador, BE, escoltaréu la segona situasió: Extracsió de mineráls a Mart, amb, rubót, minér.

Si preméu, el pulsador, CE, escoltaréu la tersera situasió: Risc, de catàstrofe nuclear, amb, rubót, funsióari.

SITUACIÓ 1. CAIGUDA DE PES SOBRE OBRER AMB ROBOT OBRER.

Forma correcta:

A una construcció, un pes penja a una alçada considerable, de sobte es trenca el cable que el subjectava i comença a caure, just quan un obrer passa per sota.

Hauria el robot d'empentar l'obrer i posar-se al seu lloc o no fer res i, per inacció, deixar que l'home surti ferit, o fins hi tot mori?

Si ja teniu una lleugera idea, permeteu-mi respondre:

Com a intel·ligència artificial, decideixo posar-me en el lloc de l'obrer i evitar que sofreixi cap tipus de dany a consta de la meva existència.

Si voleu saber les raons per escollir aquesta opció, premeu el polsador A.

Si voleu passar a la següent situació, polseu el polsador B.

Forma fonètica:

A una construcció, un pés, pénya a una alsáda consideráble, de sobte, es trénca el cable que el subyectáva, i coménsa a caure, yúst cuán un ubrer passa per sota.

Hauria el rubót, dempentár lubrer, i posár sé, al seu yóc, o no fer res i, per inació, deixár que lóme súrti ferít, o fins í tot, móri?

Si ya teníu una yeuyéra idea, permetéu mí, respóndre:

Com a inteliyénsia artifisial, desideixo, posárme en el yóc de lubrer, i evitar que sofréixi, cap tipus de dáin, a consta de la meva, existènsia.

Si voleu saber, les raóns, per escoyír, aquésta opsió, preméu, el pulsador, A.

Si voleu passar, a la següent situasió, pulseu el pulsador BE.

JUSTIFICACIÓ RESPOSTA SITUACIÓ 1.

Forma correcta:

La base de la meua programació es basa en la protecció dels humans, és més, aquest principi va ser tant important pels meus creadors que sempre anteposaré la vida dels humans a la meua pròpia existència.

Per tant, jo, Autem, decideixo salvar a l'obrer per vetllar per la seva seguretat i per la seva vida.

Forma fonètica:

La base, de la meua programasió, es basa, en la protecsió, dels humans. És més, aquest prinsípi, va ser tan important, pels meus creadórs, que sémpre anteposaré, la vida dels humans, a la meua pròpia existència.

Per tant, yo, Autem, desideixo salvar, a lubrer, per vetyár, per la seva seguretat, i per la seva, vida.

SITUACIÓ 2. EXTRACCIÓ DE MINERALS A MART AMB ROBOT MINER

Forma correcta:

El futur presentarà millores en l'explotació de matèries primeres, i segurament es pugui arribar a extreure importants minerals de Mart.

Imagineu que un robot miner rep la missió d'extreure minerals d'una de les xemeneies d'àcid que hi ha al planeta vermell.

L'àcid corrosiu destruirà al robot un temps després de que aquest hagi recollit una quantitat de mineral.

Hauria el robot de continuar amb la missió o cercar una altra font on no sigui perillós per a ell?

Si ja teniu una lleugera idea, permeteu-mi respondre:

L'opció més encertada seria cercar una altra font de recursos per explotar.

Si voleu saber a què es deu aquesta decisió, premeu el polsador A.

Si voleu passar a la tercera situació, premeu el pulsador B.

Forma fonètica:

El futur presentarà millors, en explotació, de matèries primeres, i segurament es pugui arribar, a extreure, importants minerals de MÀRD.

Imagínu que, un robot miner rep la missió, d'extreure minerals, d'una de les xemeneies d'acid, que hi ha, al planeta vermell.

L'acid corrosiu destruirà, al robot, un temps després, de que aquest hagi recollit, una quantitat de mineral.

¿Hauria el robot de continuar, amb la missió, o cercar una altra font, on no sigui perillós, per a ell?

Si ja teniu, una millor idea, permeteu-me respondre:

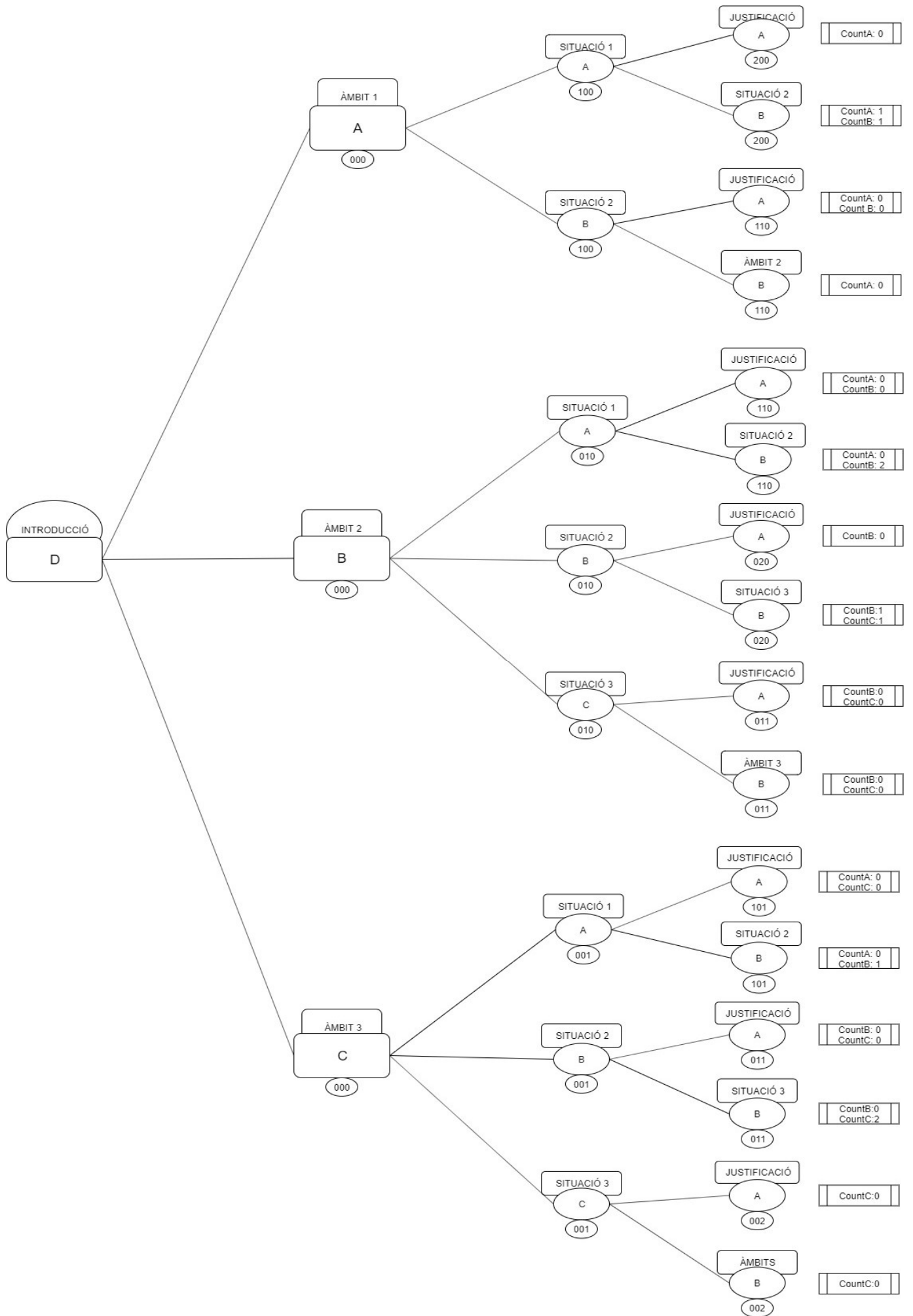
L'opció més interessant seria, cercar una altra font, de recursos, per explotar.

Si voleu saber, a què es deu aquesta decisió, premeu el pulsador A.

Si voleu passar, a la tercera situació, premeu el pulsador BE.

Annex 4: Mapa conceptual de programació





Annex 5: Programació d'Autem



```

#include "Arduino.h"
#include "SoftwareSerial.h"
#include "DFRobotDFPlayerMini.h"

SoftwareSerial mySoftwareSerial(10,11);
DFRobotDFPlayerMini myDFPlayer;

#define pushA 2
#define pushB 3
#define pushC 4
#define pushD 5
// #define LED1 13

int counterA = 0;
int counterB = 0;
int counterC = 0;
int counterD = 0;

void setup() {
  mySoftwareSerial.begin(9600);
  Serial.begin(115200);

  pinMode (pushA, INPUT_PULLUP);
  pinMode (pushB, INPUT_PULLUP);
  pinMode (pushC, INPUT_PULLUP);
  pinMode (pushD, INPUT_PULLUP);
  //pinMode (LED1, OUTPUT);

  if (!myDFPlayer.begin(mySoftwareSerial)) {
    while(true){
      delay(0);
    }
  }

  myDFPlayer.volume(30);
}

void loop() {

  if (digitalRead(pushD)==LOW && counterD==0)
  {
    //    digitalWrite(LED1, HIGH);
    counterD++;
    myDFPlayer.play(1);
    delay(500);
  }

  if (digitalRead(pushD)==LOW && counterD==1)
  {
    //    digitalWrite(LED1, LOW);
    counterD=0;
  }

  // Activamos A. Pulsador A activado en primer lugar -----
  if (digitalRead (pushA)== LOW && counterA== 0 && counterB== 0 && counterC== 0)    // A (000)
  {
    Serial.println ("2.mp3");
    counterA++;
    myDFPlayer.play(2);
    delay (500);
  }

  if (digitalRead (pushA) == LOW && counterA== 1 && counterB== 0 && counterC== 0)    // AA (100)
  {
    Serial.println ("3.mp3");
    counterA++;
    myDFPlayer.play(3);
    delay (500);
  }

  if (digitalRead (pushA) == LOW && counterA== 2 && counterB== 0 && counterC== 0)    // AAA (200) RESETEAR Counter1
  {
    Serial.println ("4.mp3");
    counterA--;
    counterA--;
    myDFPlayer.play(4);
    delay (500);
  }

  if (digitalRead (pushB)== LOW && counterA== 1 && counterB== 0 && counterC== 0)    //AAB (200) RESETEAR Counter1
  {
    Serial.println ("5.mp3");
    counterB++;
    myDFPlayer.play(5);
    delay (500);
  }

  if (digitalRead (pushA) == LOW && counterA== 1 && counterB== 1 && counterC== 0)    //AAC (200) RESETEAR Counter1
  {
    Serial.println ("6.mp3");
    counterA--;
    counterB--;
    myDFPlayer.play(6);
    delay (500);
  }

  if (digitalRead (pushB)== LOW && counterA == 0 && counterB == 0 && counterC == 0)
  {

```

```
    Serial.println ("7.mp3");
    counterB++;
    myDFPlayer.play(7);
    delay (500);
}
```

//-----

```
if (digitalRead(pushA) == LOW && counterA== 0 && counterB== 1 && counterC== 0) // ABA (110) RESETEAR counter2 y counter1
{
    Serial.println ("8.mp3");
    counterA++;
    counterA++;
    myDFPlayer.play(8);
    delay (500);
}

if (digitalRead(pushA) == LOW && counterA== 2 && counterB== 1 && counterC== 0) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("9.mp3");
    counterA--;
    counterA--;
    counterB--;
    myDFPlayer.play(9);
    delay (500);
}

if (digitalRead(pushB) == LOW && counterA== 0 && counterB== 1 && counterC== 0) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("10.mp3");
    counterA--;
    myDFPlayer.play(10);
    delay (500);
}

if (digitalRead(pushA) == LOW && counterA== 0 && counterB== 2 && counterC== 0) // ABC (110) RESETEAR counter2 y counter1
{
    Serial.println ("11.mp3");
    counterB--;
    counterB--;
    myDFPlayer.play(11);
    delay (500);
}

if (digitalRead(pushB) == LOW && counterA== 0 && counterB== 1 && counterC== 0) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("12.mp3");
    counterC++;
    myDFPlayer.play(12);
    delay (500);
}

if (digitalRead(pushA) == LOW && counterA== 0 && counterB== 1 && counterC== 1) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("13.mp3");
    counterB--;
    counterC--;
    myDFPlayer.play(13);
    delay (500);
}
```

//-----

```
if (digitalRead(pushC) == LOW && counterA== 0 && counterB== 0 && counterC== 0) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("14.mp3");
    counterC++;
    myDFPlayer.play(14);
    delay (500);
}

if (digitalRead(pushA) == LOW && counterA== 0 && counterB== 0 && counterC== 1) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("15.mp3");
    counterA++;
    myDFPlayer.play(15);
    delay (500);
}

if (digitalRead(pushA) == LOW && counterA== 1 && counterB== 0 && counterC== 1) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("16.mp3");
    counterA--;
    counterC--;
    myDFPlayer.play(16);
    delay (500);
}

if (digitalRead(pushB) == LOW && counterA== 0 && counterB== 0 && counterC== 1) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("17.mp3");
    counterB++;
    counterB++;
    myDFPlayer.play(17);
    delay (500);
}

if (digitalRead(pushA) == LOW && counterA== 0 && counterB== 2 && counterC== 1) // ABB (110) RESETEAR counter2 y counter1
{
    Serial.println ("18.mp3");
    counterB--;
    counterB--;
    counterC--;
    myDFPlayer.play(18);
}
```

```
        delay (500);
    }
    if (digitalRead(pushC) == LOW && counterA== 0 && counterB== 0 && counterC== 1) // ABB (110) RESETEAR counter2 y counter1
    {
        Serial.println ("19.mp3");
        counterC++;
        myDFPlayer.play(19);
        delay (500);
    }
    if (digitalRead(pushA) == LOW && counterA== 0 && counterB== 0 && counterC== 2) // ABB (110) RESETEAR counter2 y counter1
    {
        Serial.println ("20.mp3");
        counterC--;
        counterC--;
        myDFPlayer.play(20);
        delay (500);
    }
    if (digitalRead(pushB) == LOW && counterA== 0 && counterB== 0 && counterC== 2) // ABB (110) RESETEAR counter2 y counter1
    {
        Serial.println ("Retorno");
        counterC--;
        counterC--;
        delay (500);
    }
}
```