



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS E INGENIERÍA DE LA COMPUTACIÓN

“COMPORTAMIENTO EMOCIONAL EN ROBOTS DE  
SERVICIO”

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN CIENCIAS DE LA COMPUTACIÓN

PRESENTA:

MAURICIO ENRIQUE REYES CASTILLO

TUTOR:

DR. LUIS ALBERTO PINEDA CORTÉS  
IIMAS, UNAM

CIUDAD DE MÉXICO, JUNIO.

2018



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



# Índice general

Resumen . . . . .	1
<b>1. Introducción</b>	<b>3</b>
1.1. Emociones y robots . . . . .	3
1.2. Expresiones faciales y robots . . . . .	5
1.3. El robot de servicio Golem-III . . . . .	7
1.4. Objetivos . . . . .	9
1.5. Experimentación y evaluación . . . . .	10
<b>2. Modelos Emocionales de Computación</b>	<b>15</b>
2.1. Categorías discretas . . . . .	16
2.2. Expresiones faciales positivas y negativas . . . . .	18
2.3. Teoría de la evaluación . . . . .	21
2.4. Teoría dimensional . . . . .	22
2.5. Teorías de aproximación comunicativa . . . . .	23
2.6. Enfoque racional . . . . .	23
2.7. Teorías mixtas . . . . .	24

2.8. Sumario . . . . .	27
<b>3. Caras robóticas</b>	<b>29</b>
3.1. El valle inexplicable . . . . .	30
3.2. Diseños y aplicaciones . . . . .	32
3.3. Pruebas de identificación facial . . . . .	36
3.4. Enfoque minimalista . . . . .	37
3.5. Sumario . . . . .	41
<b>4. Robot de servicio</b>	<b>43</b>
4.1. Robots de asistencia . . . . .	44
4.2. Robots colaboradores . . . . .	45
4.3. Robots de interacción a corto plazo . . . . .	45
4.4. Robot social . . . . .	46
4.5. Robot de servicio . . . . .	49
4.6. El robot Golem-III . . . . .	52
<b>5. Prototipo facial y pruebas de identificación</b>	<b>61</b>
5.1. Prototipo facial GolemX-1 . . . . .	62
5.2. Pruebas de identificación facial . . . . .	66
5.3. Prototipo facial de Golem-III . . . . .	69
5.4. Modelo de diálogo . . . . .	76
5.5. Sumario . . . . .	78
<b>6. El efecto positivo de una retroalimentación negativa</b>	<b>83</b>
6.1. El efecto regulador de las expresiones faciales negativas . . . . .	85

6.2. Resultados . . . . .	92
6.3. Discusión . . . . .	94
<b>7. La expresión robótica de enojo en interacción humano robot (HRI)</b>	<b>97</b>
7.1. Antecedentes . . . . .	98
7.2. El efecto del enojo sobre la atención humana . . . . .	99
7.3. Cuestionario . . . . .	104
7.4. Resultados . . . . .	106
7.5. Discusión . . . . .	115
<b>8. La activación de expresiones emocionales en Golem-III</b>	<b>117</b>
8.1. Inferencia oportunista . . . . .	118
8.2. Sumario . . . . .	124
<b>9. Conclusiones</b>	<b>127</b>
<b>Bibliografía</b>	<b>137</b>
<b>A. Conceptos y desarrollo de los prototipos faciales</b>	<b>167</b>
A.1. Memoria gráfica descriptiva. . . . .	168



# Resumen

La expresión facial es un mecanismo utilizado para mostrar el comportamiento emocional durante la interacción social; sirve para mostrar la motivación e intención de los individuos, además hace entendible y predecible la conducta de las personas. Con el fin de mejorar la interacción entre humanos y robots se han implementado modelos emocionales y caras artificiales capaces de expresar emociones en sistemas autónomos; especialmente en robots sociales y de servicio. Esta investigación analiza la función de las expresiones faciales universales de mayor legibilidad generadas por un rostro robótico minimalista, diseñado y construido como instrumento de experimentación, el cual se utiliza durante una tarea colaborativa entre un humano y un robot de servicio. Los resultados de encuestas digitales, sesiones experimentales y un cuestionario, indican que la expresión de enojo es la mejor reconocida comparada con otras emociones universales, tiene un efecto regulador en las acciones e induce a la atención cuando una condición poco entendible ocurre durante una tarea. Además, se observó que la presencia prolongada de la expresión de enojo reduce su efecto sobre la atención cuando se compara con expresiones faciales positivas. Esta investigación ofrece una base para adecuar, determinar y establecer la presencia de expresiones faciales universales en un robot durante tareas de servicio. Finalmente, la activación de emociones se genera a través de modelos de diálogo y un módulo de inferencia implementado en la arquitectura cognitiva del sistema.



# Capítulo 1

## Introducción

### 1.1. Emociones y robots

El concepto general de emoción es difícil de establecer; para definirla es necesario considerar el estado mental de los individuos y la respuesta ante estímulos externos, así como las situaciones emergentes durante la interacción social. Tales respuestas repercuten en el estado físico y psicológico dando lugar a lo que conocemos como sentimientos o sensaciones. Sin embargo, la presencia de un fenómeno externo no influye de igual manera a quienes lo perciben ya que puede estar sujeta a la valoración mental y puede ser de forma consciente o inconsciente [1–8] En este caso la experiencia emocional juega un papel importante que influye directamente en la reacción ante un estímulo o la retroalimentación ante una acción previa; la interpretación a las situaciones influye decisivamente en la sensación experimentada.

Diversas teorías consideran al componente cognitivo de las emociones como el punto central para cerrar el ciclo entre estímulo y respuesta. Más allá de una definición, es posible identificar a las emociones como los medios que sirven para dar valores a las actividades cotidianas. A partir de sus manifestaciones se pueden formar interpretaciones o creencias acerca de quienes las expresan [6].

La expresión artificial de emociones ha dado nuevas formas de crear estímulos para la experimentación con humanos y robots; el papel que juegan en el comportamiento ha motivado el desarrollo de análogos computacionales de las emociones capaces de construir sistemas más inteligentes, flexibles y capaces [6], en ambientes complejos e inciertos [9]. Se ha demostrado que los sistemas robóticos expresivos favorecen a la retroalimentación y al vínculo con las personas [10–12]. Los estudios en el desarrollo de modelos computacionales se centran principalmente en el intercambio emocional. Generalmente, la evaluación de un fenómeno externo influye directamente en el cambio de emoción [6,9].

Los diseñadores de agentes emocionales de computación asignan arbitrariamente el cambio de valores en el comportamiento emocional; se ha estudiado de manera limitada los factores que afectan el cambio de estado así como el impacto de las emociones en los usuarios. La investigación de ambos elementos reforzará el diseño de sistemas que se acoplen a las expectativas del usuario para lograr que la interacción sea adecuada al contexto y a la situación.

Los estudios en interacción humano robot han tomado en cuenta algu-

nas señales sociales tales como las expresiones faciales, el lenguaje corporal y el diálogo [10, 18, 19]. Para estudiar su influencia en las personas se ha evaluado el tipo de reacción y el tiempo de respuesta. Durante la evaluación del desempeño del sistema se toma en cuenta la eficiencia en la tarea y la expectativa del usuario respecto a las acciones e intenciones del robot.

## 1.2. Expresiones faciales y robots

Las expresiones faciales son un mecanismo común para evidenciar el comportamiento emocional durante la interacción social; sirven para mostrar motivación individual y hacer al comportamiento tanto entendible como predecible. Además, son útiles para sincronizar, organizar y finalizar la interacción [13]. En robots interactivos es común encontrar diseños de prototipos faciales; algunos de los rostros robóticos capaces de ejecutar gestos requieren de sistemas complejos electromecánicos los que consideran a los 25 músculos que actúan en el rostro humano para tal función. Existen estudios que categorizan los tipos de rostros a partir de los grados de libertad implícitos (DOF) y el aspecto, que va de lo icónico a lo realista [12, 14, 15]. Para facilitar el estudio de rostros robóticos y el uso de expresiones faciales en la interacción humano robot se han generado diversos sistemas que minimizan el equipo y material de construcción con el máximo de efectividad.

Dada la diversidad en el diseño de sistemas faciales, la interpretación de la emoción es variable y suele estar ligada al criterio cultural, pese a su

universalidad. Las expresiones faciales están presentes durante la interacción social y el contexto afecta su sentido de muchas maneras. Un problema en el estudio de las expresiones faciales es la diversidad de interpretaciones en culturas diferentes [10]; con el fin de reducirlo, diversos estudios se centran en el conjunto de expresiones faciales básicas o universales [14, 16]. Ekman propone los siguientes: felicidad, sorpresa, miedo, disgusto, tristeza y enojo [3, 17], los cuales se han definido como biológicamente universales en la especie humana a partir de su reconocimiento y aparición en distintas culturas. Su representación en diversos sistemas robóticos se ha evaluado en pruebas para la identificación facial. Durante dicha evaluación se presentan las imágenes o el prototipo físico del sistema a una o a varias personas quienes asignan una etiqueta emocional al tiempo que se ejecuta una expresión en particular. Gracias a estas evaluaciones se puede dar dirección al diseño y programación de una cara robótica útil durante la interacción con personas.

Los modelos de emoción integran categorías discretas basadas en las expresiones básicas universales. Para explicar la dinámica, transición e intercambio de las emociones se utilizan escalas continuas representadas en espacios de dos o tres dimensiones [17]; en el caso de los planos tridimensionales (plano interpolador tridimensional) se asignan valores como la excitación (alta o baja), la valencia (negativo o positivo) y la postura (abierta o cerrada) [4, 5, 9]. Es importante considerar los factores que influyen en el cambio de valores y el estado emocional; por ejemplo, de felicidad a tristeza, de tristeza a enojo, etc.

Por lo regular, la evaluación directa de la expresión robótica sobre las acciones humanas se ha probado de forma limitada. Los análisis desarrollados, se han probado como complemento de otras modalidades del comportamiento emocional (expresión facial en conjunto al diálogo, expresión corporal, entre otras). La información respecto a la influencia directa de las expresiones faciales en humanos a través de un robot de servicio es aún limitada; dada las características y funciones de las emociones a nivel social resulta un campo nuevo a la investigación. El fenómeno resulta de mayor interés si se integra en un entorno de servicio y colaboración con el fin de favorecer la interacción entre usuario y sistema.

Esta investigación inédita, presenta el desarrollo de criterios para la activación emocional artificial tomando en cuenta la influencia de la expresión facial robótica en los humanos y dentro de un entorno colaborativo. Los análisis tienen el fin de adecuar, determinar, establecer y activar oportunamente las expresiones faciales artificiales a través de un comportamiento emocional integrado a la arquitectura cognitiva de un robot de servicio.

### **1.3. El robot de servicio Golem-III**

Golem-III es un robot humanoide de servicio desarrollado en el Instituto de Investigaciones en Matemáticas Aplicadas y Sistemas (IIMAS) y fabricado parcialmente en el Centro de Investigaciones de Diseño Industrial (CIDI) de la UNAM. El proyecto se desarrolla por el Grupo Golem (IIMAS) y consiste en la integración de una estructura móvil capaz de

ejecutar el registro y emisión de audio, la manipulación de objetos, navegación, el reconocimiento facial y corporal, la captura de imágenes y video, el reconocimiento de objetos, entre otros comportamientos.

El objetivo principal del proyecto Golem se centra en el desarrollo e implementación de un modelo conceptual de un robot de servicio; sistemas que desempeñan tareas específicas para los humanos. La programación y el comportamiento del sistema están diseñadas para interactuar con personas dentro de un entorno susceptible a variables [20–22]. Básicamente, la arquitectura robótica de Golem-III le permite adaptarse a los estados emergentes en el entorno y cumplir con sus tareas. Su diseño se define por la relación existente entre las entradas y salidas del sistema. Para efectos de esta investigación se desarrollaron varios prototipos faciales electromecánicos para Golem-III los cuales se utilizaron de manera independiente y bajo condiciones controladas. Un tema de interés para el Grupo Golem es la expresión emocional en el robot de servicio.

Para integrar un sistema periférico en Golem-III capaz de mostrar un comportamiento emocional, se diseñó un prototipo facial que utiliza el mínimo de elementos maximizando la eficiencia durante la ejecución de expresiones básicas. El rostro se realizó de forma digital previa a su fabricación. Se diseñó una encuesta en línea en la que se presentaron a varias personas una serie de imágenes del rostro robótico ejecutando expresiones universales; los sujetos nunca habían visto la propuesta con anterioridad. De forma similar a los métodos para evaluar la legibilidad de la expresión, el diseño se mostró al azar ejecutando 5 de las 6 emociones universales (felicidad,

tristeza, enojo, sorpresa y miedo). Además, se realizaron algunos experimentos para analizar la influencia de las expresiones en los humanos en el contexto de una tarea simple bajo condiciones controladas. Por último, se desarrollaron varios prototipos faciales electromecánicos que sirvieron de base para la fabricación de un sistema facial óptimo integrado en el Robot Golem-III.

## 1.4. Objetivos

Este trabajo tiene como objetivo general establecer las bases para la integración de un comportamiento emocional en un robot de servicio que influya positivamente en el comportamiento del usuario.

El objetivo general se concreta a través de los siguientes objetivos particulares:

- Analizar el desarrollo de modelos emocionales artificiales para robots interactivos.
- Identificar los elementos más utilizados para la expresión emocional en robots interactivos.
- Desarrollar un sistema para el estudio de la expresión de emociones durante la interacción humano y robot.
- Validar el efecto de las emociones robóticas en el usuario para determinar y adecuar su activación.

- Implementar el comportamiento emocional en un robot de servicio durante tareas e interacción con el usuario.

## 1.5. Experimentación y evaluación

Con el fin de llevar a cabo los objetivos planteados se diseñó, evaluó y desarrolló un rostro robótico capaz activar expresiones faciales a partir del estudio de las emociones universales y los modelos computacionales de emoción. Se realizaron pruebas de expresión facial con el fin de optimizar el prototipo. Durante el análisis, se prestó atención a los gestos de mayor legibilidad con el fin de obtener un efecto positivo en el comportamiento humano en relación a una tarea del robot de servicio. La mejor legibilidad fue registrada por las expresiones de tristeza y enojo, lo cuál es consistente con las teorías psicológicas; se argumenta que la mejor identificación de las expresiones negativas es debido a la relación existente entre su activación y la presencia de un fenómeno adverso o dañino en el entorno [23–26].

Las emociones negativas se ligan al aspecto evolutivo de los seres humanos. Debido a esto, se decidió centrar esta investigación en la influencia de las expresiones negativas y su contribución durante la interacción humano robot, principalmente en su efecto positivo en tareas de un robot de servicio. La experimentación con una actividad de interacción humano y robot (uno a uno) ha mostrado ser efectiva para analizar los efectos de las acciones realizadas por el robot; específicamente, señales sociales como el seguimiento ocular, el dialogo, las expresiones emocionales en rostro y

cuerpo [27, 28]. La implementación del comportamiento social en robots interactivos ha resultado benéfica para su presencia en ambientes sociales [29, 30, 101]; sin embargo es un tema sumamente complejo abierto al análisis.

El prototipo facial desarrollado se adaptó para interactuar directamente con el usuario; para ello se fabricó un torso y brazos con el fin de implementar un escenario de experimentación en simulación a una tarea colaborativa. Se consideró este tipo de interacción con base a la importancia de la participación humana en la actividad y con el fin de analizar los efectos directos del gesto en el comportamiento. La tarea era simple para evitar la intervención prolongada del sujeto y sobre todo la ejecución de maniobras complejas de parte del humano y el sistema.

Los experimentos permitieron observar el efecto positivo que resulta de la ejecución de expresiones negativas de tristeza y enojo. Cabe señalar que el enojo influyó en mayor medida sobre las acciones y atención humana. De acuerdo a la literatura en psicología, el enojo ha sido sujeto a cambios en su efecto debido a la adaptación en los entornos sociales y tiene un impacto significativo en las relaciones interpersonales de negociación y de beneficio en quien lo expresa [23, 32, 33].

En la primera sesión experimental se diseñó una tarea en la que humano y robot colaboraron. Se estudió el efecto de las expresiones robóticas sobre las actividades del sujeto; se observaron las acciones humanas que ayudaban al progreso de la tarea sin omitir las reglas estipuladas de antemano. La activación del rostro se asignó a las situaciones en la interacción, en este

caso, el de felicidad en relación al progreso de la tarea y el de tristeza y enojo durante una inconsistencia generada por una falla (programada) en el robot. La presencia del fallo y la activación emocional dieron la pauta para la intervención del sujeto. De esta forma, la variable independiente fue el rostro activado durante el fallo y la dependiente al tipo de respuesta y acción del humano. El rostro neutral en el robot se utilizó como control durante una sesión completa.

El resultado evidenció que el uso de expresiones faciales puede regular el comportamiento de la persona; su presencia indica una situación adversa al progreso de la tarea en la cual el sujeto se percata de la activación facial, asume la condición de fallo y contribuye al flujo de la actividad. A diferencia de la omisión de emoción (neutral) la emoción negativa mostró tener una mayor influencia positiva para completar la interacción. Es importante señalar que la expresión de rostros negativos puede tener un impacto adverso en sujetos humanos, por lo que en esta investigación se rescatan los aspectos positivos durante su activación. Estadísticamente no hubo diferencia entre la expresión de tristeza y enojo cuando se presentaron al usuario; las respuestas del sujeto ante las emociones fueron similares, excepto por la diferencia en el tiempo de reacción favorable al enojo (el sujeto reaccionó más rápido).

En el segundo experimento se analizó el efecto del enojo sobre la atención humana; se eligió el rostro que obtuvo los mejores registros en las evaluaciones anteriores además de sus cualidades de acuerdo a la literatura en psicología. Se diseñó una tarea en la cual se utilizaron las mismas

reglas que en el experimento anterior; sin embargo, existió una condición poco evidente la cual activaba la emoción; la expresión facial era diferente a los presentes durante el progreso de la tarea. El fin del experimento era determinar si el enojo inducía al sujeto a analizar y determinar la causa de la expresión. La variable independiente fue el rostro activado durante la situación indefinida y la variable dependiente fue la respuesta del colaborador respecto a las supuestas causas que activaron las emociones; se aplicó una encuesta al final de la interacción individual. Durante las sesiones se compararon las expresiones negativas y positivas (enojo contra felicidad y sorpresa).

El resultado mostró que el enojo induce a la atención por encima de las emociones positivas; los sujetos identificaron con mayor precisión las causas en la activación de expresiones. Sin embargo, el enojo tiene un mejor resultado cuando es activado de forma súbita a diferencia de su presencia continua e incremental; se implementó el estado dinámico de las expresiones a través de su transición neutral a enojo con alta intensidad.

Los estudios realizados permiten identificar las cualidades positivas del enojo para su activación y uso en comportamientos artificiales de emoción. El manejo adecuado de la emoción puede influir favorablemente a la interacción con humanos durante actividades colaborativas. La generación del enojo dentro de las conductas emocionales para robots sociales o de servicio puede representar un riesgo ante las posibles respuestas de los usuarios, tener un efecto adverso e incluso carecer de objetividad dentro de la interacción. Sin embargo, anticipar y deducir las reacciones posibles de los

usuarios ante la emoción permite establecer la objetividad en el uso de expresiones artificiales a través del uso de dispositivos y sistemas capaces de representar comportamientos emocionales.

## Capítulo 2

# Modelos Emocionales de Computación

Las señales sociales no verbales son útiles durante la interacción entre robots y personas en ambientes reales. Los sistemas que las utilizan son capaces de señalar, observar, expresarse con el cuerpo y la cara. Comúnmente se conocen como robots de interfaz social a los sistemas que utilizan medios empleadas en la comunicación humana [34]. Para explicar el proceso que permite la emisión de emociones se ha sugerido la interacción de dos sistemas; el cognitivo y el emocional. El primero es el responsable de interpretar los fenómenos del mundo real, mientras que el segundo evalúa y juzga el valor general de tales fenómenos. Cuando ambos sistemas operan adecuadamente, el sistema emocional modula los parámetros del sistema cognitivo y da una nueva interpretación de lo observado [35].

Inicialmente las emociones se interpretaron en términos de categorías discretas (ej. Felicidad, Tristeza, etc.). Ekman ofrece una revisión de emociones reconocidas de manera transcultural (felicidad, tristeza, enojo, disgusto, sorpresa y miedo) que denominó como básicas universales; en sus estudios encontró que los miembros de una cultura aislada (tribu de Papúa Nueva Guinea) fueron capaces de identificar fácilmente expresiones emocionales al observar fotografías tomadas a personas de culturas diferentes. Los sujetos también adjudicaban situaciones específicas a ciertas expresiones con un criterio similar al de otras culturas [36, 37].

Más adelante se utilizaron escalas continuas con dos valores representados como ejes dentro de una gráfica bidimensional: la excitación/activación, (arousal) y la valencia, (valence) [38]. El concepto más explorado se compone de la importancia de las dos anteriores; categorías y dimensiones. Ambas en estado dinámico dentro de un espacio sujeto a variantes del exterior [1, 5, 39, 40]. Las teorías que han demostrado tener mayor impacto en sistemas computacionales argumentan la evaluación de eventos y la reacción subsecuente en un ambiente determinado [6, 41–44]. A continuación, se revisan algunas teorías utilizadas en modelos computacionales de emoción.

## 2.1. Categorías discretas

Las expresiones faciales y emociones básicas han sido categorizados para explicar la emisión, intercambio de estados y su respuesta a eventos en el entorno. Algunos modelos de evaluación computacional utilizan el conjunto

de variables para el procesamiento y expresión de emociones propuestas por Ortony, Clore y Collins [4, 45]. En su estudio se proponen 22 emociones basadas en categorías a partir de la evaluación a situaciones y eventos específicos. Sin embargo, reducir la cantidad de expresiones con el fin de lograr una interpretación universal se ha investigado a profundidad [46]. De acuerdo con Ekman, el reconocimiento de expresiones faciales en diversas culturas establece seis emociones: Alegría, Tristeza, Enojo, Miedo, Sorpresa y Disgusto (Figura 2.1 [47]). La activación las expresiones faciales tienen una función social predeterminada a partir de ciertas condiciones que la detonan y funciones en el entorno de interacción; en el cuadro 2.1 se muestra tal relación en los gestos universales [2, 9].

Las emociones universales han servido para explorar modelos computacionales utilizados durante la interacción entre humano y computadora (HCI) [48] e interacción humano robot (HRI) [10, 15, 18]. Actualmente las aplicaciones computacionales y robóticas son capaces de configurar casi todas las expresiones básicas y otras no consideradas dentro del criterio universal.

La síntesis en la representación de expresiones faciales ha tenido una amplia difusión en redes sociales a través de la secuencia de caracteres denominados emoticonos, estos se aplican en robots con rostros minimalistas. La síntesis en los gestos se utiliza para pruebas de interacción humano-robot sin la necesidad de invertir en el desarrollo de rostros complejos y caros [14, 49].

Cuadro 2.1: Condiciones y funciones de las emociones universales [2].

Condición	Emoción	Función
Ausencia de un estímulo o expectativa deseada	<b>Tristeza</b>	Evocar comprensión y solidaridad
Presencia de un estímulo o su extensión	<b>Calma</b>	Continuar la interacción con el estímulo deseado
Éxito o logro de un objetivo	<b>Felicidad</b>	Expectativa positiva por aun nuevo estímulo
Dificultad para conseguir algo. Un comportamiento adaptativo	<b>Enojo</b>	Incomodidad. Influir en el cambio de una situación
Presencia de un estímulo indeseable	<b>Disgusto</b>	Señal que un estímulo no deseado sigue presente en el entorno
Una situación repentina. Un estímulo agradable no esperado	<b>Sorpresa</b>	Satisfacción repentina. Advertir de una situación
Ansiedad ante un estímulo amenazante	<b>Miedo</b>	Reaccionar repentinamente. Alejarse del estímulo

## 2.2. Expresiones faciales positivas y negativas

Las emociones se dividen ocasionalmente en categorías negativas y positivas, tiene su origen en el campo de la ética y otras áreas que enfatizan esta dicotomía; por ejemplo la valencia en Física y la polaridad de la Química [23, 24, 37, 46]. De manera alternativa, la excitación y la valencia son propuestas como dimensiones extra para distinguir el estado emocional. Estas



Figura 2.1: Ilustraciones con las seis expresiones universales del proyecto “Grimace” (Desde arriba Izquierda) Enojo, Felicidad, Sorpresa, Disgusto, Tristeza, Miedo [47].

son representadas en el Modelo Circumplejo (Circumplex Model of affect) [5,50–52]. Por ejemplo; el enojo es una emoción negativa de excitación alta y del lado contrario es una positiva de baja excitación la cuál se representa por el estado contento o de modo relajado. De forma similar, en el modelo de Activación positiva - Activación negativa (positive activation - negative activation model. PA-NA) [53–55], enojo es de activación negativa alta y en su lado opuesto se encuentra una de activación positiva baja (calma) (Figura 2.2).

Las expresiones emocionales influyen en el comportamiento y la atención, particularmente, las emociones negativas tienen mayor impacto sobre las positivas (Tristeza y Felicidad) ya sea comparadas de forma evidente [56] o indirecta (no deliberada) [26]. La teoría que sustenta el reconocimien-

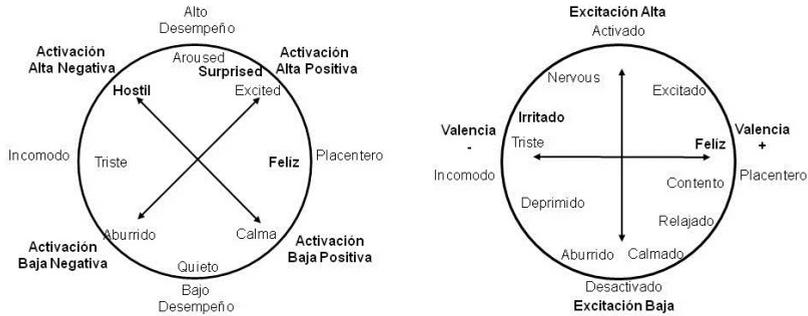


Figura 2.2: Modelos dimensionales de emoción. (Izquierda) Modelo de Activación positiva - Activación negativa (PA-NA) [53,54]. (Derecha) Modelo Circumplejo de Afecto (Circumplex Model of Affect) [5,50–52]. Las figuras muestran similitudes en las emociones y su contraparte (esta comparación es presentada en: Larsen, J., MacGraw, A. 2011) [55]

to de rostros negativos sobre los positivos tiene sus bases en la adaptación biológica y es útil para transmitir una amenaza potencial a los miembros de grupos sociales [25].

El uso de señales sociales negativas utilizadas como retroalimentación en robots puede tener efectos positivos en tareas con humanos. Este efecto ha sido probado en sistemas artificiales para influenciar al usuario o colaborador [11,57]. A continuación, se resumen las teorías sobre los modelos computacionales de emoción.

## 2.3. Teoría de la evaluación

La perspectiva psicológica predominante sobre las emociones es la teoría de la evaluación, ha servido como base para el diseño de sistemas de inteligencia artificial. Esta teoría enfatiza y explica la conexión entre la emoción y la cognición; se argumenta la aparición de los patrones de juicio individual sobre la relación entre eventos, creencias, deseos e intenciones en relación al entorno [58,59]. Gracias a esta relación se orientan las preferencias a determinados eventos mediante estructuras de significado situacional o variables de evaluación [42].

Los modelos se asocian con reacciones de la conducta, de tal forma que las evaluaciones provocan respuestas cognitivas conocidas como estrategias de afrontamiento [6, 60]. Las investigaciones se centran en la relación estructural entre las variables de evaluación y las etiquetas emocionales; se definen patrones y se traducen en alguna emoción correspondiente.

Los modelos de evaluación en el comportamiento computacional se derivan en planes teóricos, acciones de reacción, procesos de decisión de tipo Markov y modelos cognitivos detallados. La valoración de la situación se considera como la causa de la emoción a partir de una clasificación simple tal como una etiqueta (tristeza, enojo, complacencia, etc.) y se derivan con reglas si-entonces (If - Then) sobre el conjunto de variables [6].

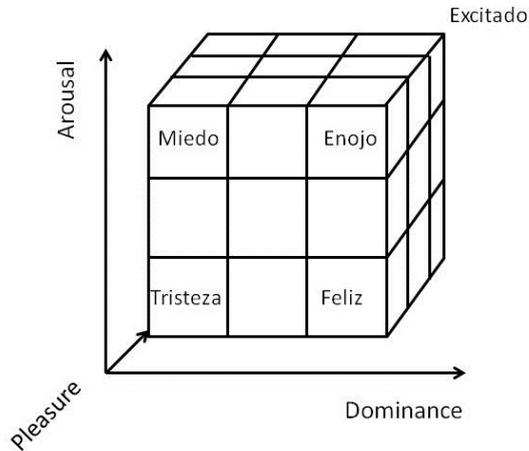


Figura 2.3: Escala de interposición tridimensional (PAD: Arousal, Pleasure y Dominance)

## 2.4. Teoría dimensional

La teoría dimensional difiere de la anterior en que se considera a los estados emocionales como variables continuas incrementales o decrementales y no como entidades discretas. Generalmente se representan en planos de dos a tres dimensiones [61] y se asignan conceptos como: placentero, activación y control. Tales valores son representados en ejes de una gráfica de coordenadas denominada escala de interposición tridimensional (PAD: Pleasure, Arousal y Dominance) (Figura 2.3). A partir de un punto central, los valores permanecen en estado dinámico resultando en una condición emocional que puede ser diferente en cada individuo sin importar que los estímulos y condiciones en el entorno sean similares.

Esta teoría resta importancia al término emoción y enfatiza el uso de tributos cognitivos, así como al balance activo entre deseos, intenciones y creencias asignando un estado individual único; personalidad y temperamento [6]. Los modelos de interposición tridimensional describen el comportamiento en tres dimensiones de forma continua y lineal mientras que en la teoría de la evaluación los comportamientos se asocian a un amplio número de valores de apreciación.

## **2.5. Teorías de aproximación comunicativa**

Esta teoría se desarrolló para influir directamente en el usuario, específicamente, durante la ejecución de tareas en las que se muestra un estado interno del sistema.

En este enfoque se enfatiza la comunicación social rompiendo la asociación entre el procesamiento interno cognitivo y expresión emocional; la emoción mostrada no necesita tener como base un estado interno pues se activa un gesto capaz de generar un efecto deseable en el usuario [62]. Los modelos de este tipo se utilizan en sistemas pedagógicos durante actividades dicotómicas de tipo recompensa-castigo o éxito-fracaso.

## **2.6. Enfoque racional**

El enfoque racional se crea a partir de la capacidad humana de adaptación emocional a condiciones en el entorno social, su utilidad durante la interacción fue la base para interpretar un modelo de inteligencia artificial

capaz de generar una conducta adaptativa [63–67]. En esta teoría, la cognición es analizada como una colección de procesos simbólicos útiles para funciones específicas e implementadas en arquitecturas limitadas al entorno de uso; la emoción es generada a partir de procesos y restricciones con valor adaptativo. Los modelos computacionales generados bajo este criterio están dirigidos para mejorar sistemas inteligentes.

## 2.7. Teorías mixtas

La teoría de la evaluación domina por mucho la aplicación de modelos computacionales, sin embargo, se han propuesto teorías compuestas para expresar emociones. La fusión entre la de evaluación y la dimensional es recurrente en arquitecturas robóticas. Se asimila a un espacio dimensional de emociones discretas en estado dinámico. Un factor importante en la teoría mixta, es la coordinación entre factores de evaluación y un estado emocional resultante, en otras palabras, la consecuencia de la emoción y la asignación de atributos emocionales a una variante externa. En este sentido, se puede dar una representación dimensional (placer, valencia y excitación) con una estructura jerárquica de emociones discretas. Un ejemplo es el modelo desarrollado por Jonathan Gratch y Stacy Marsella (EMA: A computational model of appraisal dynamics) cuyo proceso mantiene evaluaciones continuas y permite valores dimensionales de estados de ánimo que moderan a su vez la respuesta de emociones discretas [68–71].

Breazeal [9] elaboró un modelo emocional híbrido con características si-

milares y es utilizado en robots sociales; cuenta con tres dimensiones tales como la excitación, la valencia y la postura (Figura 2.4). La arquitectura cognitiva que utiliza el modelo de Breazeal, se basa en el registro de estímulos externos; un sistema de percepción activa una serie de códigos que influyen sobre el estado interno del sistema y su respuesta se genera a través de una conducta. El resultado es un conjunto valores que se evalúan en tiempo real; se asignan números enteros de forma subjetiva a la calidad y formato del estímulo (tiempo de exposición, características físicas de un objeto, intensidad y tono de voz del usuario que interactúa con el sistema). El valor (negativo o positivo) determina la posición de una etiqueta emocional dentro del plano tridimensional; la excitación (baja o alta), la valencia (positiva o negativa) y la postura (abierta o cerrada). El comportamiento del sistema se muestra a través rostro robótico y es el resultado del intercambio entre el sistema de percepción y el de activación (Proceso sensor-motor). El sistema puede cambiar el valor que le da a los estímulos dependiendo de su exposición a estos.

Marsella, Gratch y Petta proponen una arquitectura de evaluación que considera varias teorías (Figura 2.5). En el diagrama, la información fluye en un ciclo cerrado e inicia con la relación persona - ambiente (influencia sobre los hechos externos, reales o hipotéticos de una persona) y conduce a la respuesta afectiva con cierta intensidad y variables; la reacción e intensidad correspondiente al patrón de evaluaciones. La respuesta se traduce en una conducta o comportamiento; hay un cambio en la relación persona/ambiente respecto al estado original. Las etapas son referencias

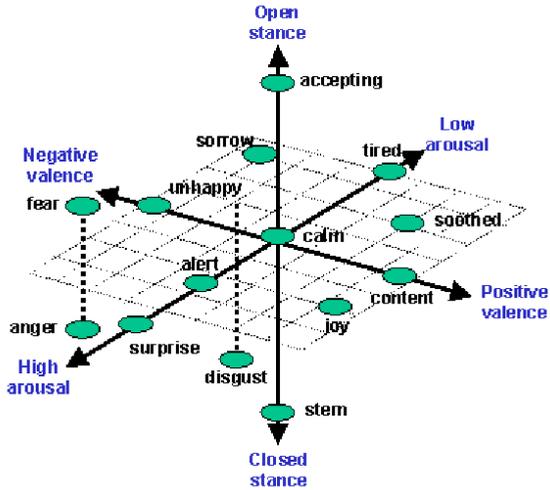


Figura 2.4: Interpolación tridimensional de etiquetas emocionales (Affect space) [9]

utilizadas por otros teóricos, cada una de estas puede ser representada por modelos que describen el proceso de la información para la ejecución emocional.

El esquema presenta un ciclo cerrado que explica la regulación autónoma de emociones en agentes artificiales. A partir de la relación persona-entorno, las variables externas son evaluadas a través de un proceso cognitivo afectando el estado físico y mental del observador lo cual repercute en el comportamiento, así como en la re-interpretación del entorno. Este ciclo cerrado pone en práctica continua la evaluación, respuesta y re-evaluación de los fenómenos externos; los agentes autónomos son capaces de adaptarse y evolucionar en consecuencia.

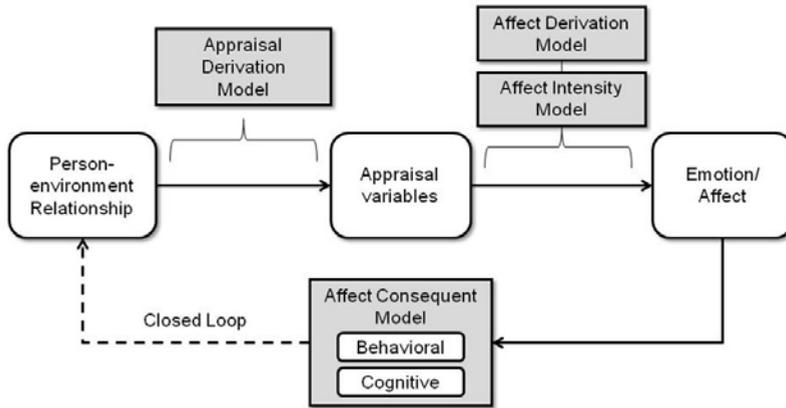


Figura 2.5: Diagrama del modelo computacional de la evaluación [6]

La terminología en esta arquitectura está asociada a las investigaciones de Lazarus (Relación persona-ambiente) [59], Gebhard y Mariner (Modelo de evaluación-derivación) [72, 73], Ortony, Clore, Collins, Elliot, Becker-Asano y Wachsmuth (Variables de evaluación) [4, 75], Ortony, Clore, Collins, Elliot, Becker-Asano, Gratch y Marsella (Modelo de afectación - derivación) [4, 69, 70, 74, 76], Gratch, Marsella y Reisenzein, (Modelo de afectación-intensidad) [69, 70, 77], Neal Reilly, Becker-Asano y Wachsmuth (Emoción-afectación) [75, 78], Becker-Asano, Gratch, Marsella, Bui, Gmytrasiewicz y Lisetti (Modelo de afectación-consecuencia) [69, 70, 76, 79, 80].

## 2.8. Sumario

Los modelos de emoción permiten a los sistemas utilizar las variables externas presentes en el entorno (reales o hipotéticas) para procesarlas a

través de un sistema cognitivo y tener una respuesta emocional; esta puede ser representada por expresiones faciales en el caso de señales sociales no verbales. A partir de los estados emocionales dinámicos se genera una interpretación del estado mental del sistema, el cuál puede estar condicionado a una actividad, variables adversas o favorables a la ejecución de acciones, logro de objetivos y postura hacia una situación. Un agente artificial con tales características puede interactuar en un entorno social de forma adecuada. El modelo de emoción ideal está compuesto por diversas teorías que permiten combinar sus elementos más representativos; categorías discretas, evaluación y respuesta, estados dinámicos de la emoción, conductas adaptativas en entornos limitados, así como respuestas emocionales capaces de generar efectos deseables en el usuario. Este último elemento ha sido explorado de forma moderada; su análisis puede contribuir al diseño de arquitecturas emocionales que favorezcan a la comunicación humano-robot, específicamente cuando el usuario desempeñe un papel clave dentro de la interacción.

## Capítulo 3

# Caras robóticas

Los robots interactivos utilizan señales sociales para comunicar su estado “mental” y “emocional”. La expresión facial es uno de los canales más utilizados y sirve para generar un comportamiento emocional inteligente. Las expresiones se activan con los estímulos del entorno, durante el logro de objetivos, al tomar decisiones, para iniciar o finalizar una interacción social [6, 10, 18].

Las expresiones faciales evocan respuestas en el observador; la más común es imitar la emoción que se observa. Sin embargo, la réplica a la expresión es diferente o nula cuando se realiza por una cara robótica sin importar que sea plenamente identificada [81]. La manera en cómo los gestos artificiales influyen en los humanos es aún una pregunta abierta.

Desde hace más de treinta años se han desarrollado caras robóticas para su estudio durante la interacción social [10]. La dimensión y rango de los diseños es muy amplio; desde caras con rasgos animales hasta rostros

abstractos y humanoides de alto realismo (androides). El reconocimiento de las expresiones es un tema importante durante la generación de prototipos; se argumenta que las caras abstractas tienen una mayor legibilidad en las expresiones [82]. Replicar las características y efectos de un rostro humano representa un reto para los desarrolladores.

### 3.1. El valle inexplicable

Los diseños de sistemas robóticos se han perfeccionado en su software y hardware; su comportamiento aunado a su apariencia les permite imitar a los humanos de forma limitada. Por un lado, las arquitecturas desarrolladas son capaces de generar sistemas reactivos al entorno y por el otro, algunos robots tienen apariencia humana gracias a estructuras basadas en el esqueleto y materiales plásticos semejantes a la piel. Cuando los agentes artificiales tienen diseños y comportamientos con rasgos antropomorfos generan expectativas positivas [10, 18]; son percibidos con naturalidad en el entorno (familiaridad). Este concepto se puede representar con una función matemática  $y = f(x)$ , en donde el valor de  $y$  aumenta (o disminuye) continuamente con el valor de  $x$ ; en este caso, un diseño cada vez más parecido a los humanos incrementa sus expectativas. Sin embargo, se observó que esta relación continúa hasta que la familiaridad disminuye abruptamente; sin importar que el parecido y el comportamiento sigan aumentando [83]. La gráfica que representa esta tendencia muestra un valle en la función que vuelve a incrementarse en la familiaridad cuando el parecido al humano es

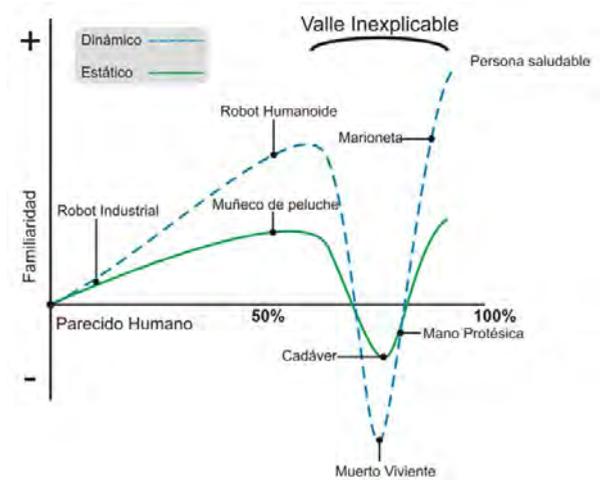


Figura 3.1: Diagrama del “Valle inexplicable” (“uncanny valley”) de Masahiro Mori. Se comparan agentes artificiales estáticos y dinámicos [85]

sumamente alto; Masahiro Mori llamo a este valle “uncanny valley” (Figura 3.1) (valle extraño/inexplicable).

Mori argumentó tal fenómeno basándose en un criterio generalizado acerca del diseño y a la ejecución de movimientos en robots; consideraba que cuando se trata de desarrollar sistemas artificiales similares a los humanos existe una condición que genera una familiaridad negativa. El valle que se genera en la función es producto del decremento en las expectativas; lo cual supone una aversión hacia el robot. En la gráfica se consideran agentes y objetos con rasgos humanos en estado estático y dinámico; desde muñecos de peluche, marionetas, prótesis humanas, robots industriales hasta robots humanoides. Tal rechazo hacia el agente es provocado por su apariencia similar a un cadáver (estado estático) o a un muerto viviente

(estado dinámico). Este concepto ha prevalecido en el desarrollo de robots humanoides y es considerado durante el diseño de sistemas socialmente interactivos. La teoría de Mori es controvertida debido a que no se realizaron pruebas empíricas para su fundamento [87]

La teoría del “Valle inexplicable” se utiliza para el diseño de rostros robóticos y la activación de expresiones emocionales; es posible hacer una categorización con el fin de identificar las dimensiones y aceptación de caras artificiales [83,84]. Estudios recientes demuestran que los rostros mecánicos y realistas tienen una influencia positiva en la familiaridad la cuál influye en el desempeño durante la interacción humano y robot. [9, 12, 14, 85, 86]. Las personas perciben favorablemente a los diseños con apariencia mecánica-humanoide por encima de los realistas; los gestos faciales generan más confianza. Estos análisis ubican al estudio de HRI directamente en las áreas de la psicología humana; los robots tienen un impacto directo en los factores socio psicológicos y afectivos [85].

## **3.2. Diseños y aplicaciones**

Hay muchos ejemplos de sistemas con inteligencia artificial que utilizan expresiones para comunicar emociones [9, 12, 14, 85, 88–91, 95]. Algunos de los diseños tienen desarrollos mecánico-eléctrico e industrial [92], con aplicaciones faciales digitales [88, 89] o sistemas híbridos electromecánicos y digitales [93]. Generalmente se postulan dos dimensiones para el desarrollo electromecánico de un prototipo facial: los equipos que imitan la cara hu-

mana e incluso utilizan motores en relación a los músculos del rostro [12,96] y los que utilizan el mínimo de componentes [14]. Se espera que ambos tipos de sistemas sean capaces de representar expresiones faciales de forma legible.

Algunos ejemplos de sistemas que expresan emociones básicas universales son el robot Kismet [9,94], los desarrollos robóticos, WE-3RIV, WE-4RII y Kobian-R [15,90,91,97–100], los robots humanoides (androides) Actroid-SIT [12], Geminoid-F [95] y ROMAN [96], el robot Flobi [92], I-Cat [16], Robotinho [101], Nexi [102], IRL-1 (Robot Melvin) [103], el robot híbrido electromecánico y digital BERT [93], Kaspar [104] que se utiliza para interactuar con niños autistas y el minimalista MiRAE [14]. Algunos otros equipos están diseñados para llevar a cabo actividades al servicio de las personas y establecer una interacción social, tales como el robot Lisa de la Universidad de Koblenz-Landau en Alemania [88] y Bender de la Universidad Católica de Chile [105]. En la figura 3.2 se muestran los rostros analizados (de mayor a menor uso de rasgos humanos). La expresión de emociones activadas durante tareas de robots interactivos de servicio es un tema en desarrollo.

Regularmente, los rostros robóticos aplican el criterio emocional y cognitivo para activar expresiones universales; sin embargo, algunos sistemas lo hacen a partir de un módulo situacional, el cuál organiza los gestos de forma adaptativa y totalmente restringida al entorno. Este método está basado en el enfoque racional expuesto en el capítulo anterior. En la investigación de Kondo y Takemura, (Actroid-SIT [12]) se desarrolló un sistema

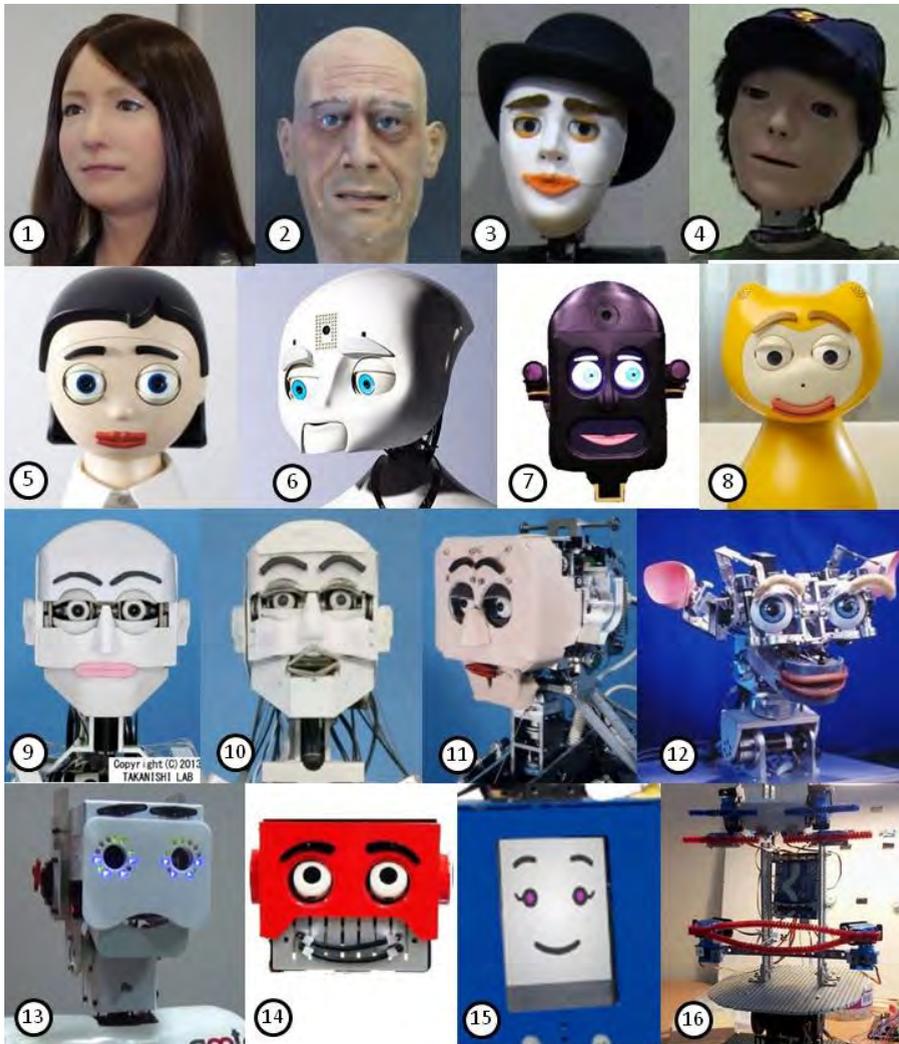


Figura 3.2: Prototipos faciales expresivos. 1.- Geminoid-F [95], 2.- RO-MAN [96], 3.- Robotinho [101], 4.- Kaspar [104], 5.- Flobi [92], 6.- Ne-xi [102], 7.- BERT [93], 8.- iCat [16], 9.- Kobian-RIV [15,91], 10.- WE-3RIV [97], 11.- WE-4RII 4 [98], 12.- Kismet [9], 13.- Bender [105], 14.- IRL-1 Melvin [103], 15.- Lisa [88], 16.- MiRAE [14].

basado en la administración de movimientos nombrado “Reconfiguration Motion Database”, cuya función es la planificación de acciones del robot. Esta estrategia reduce el tiempo de espera durante la comunicación con humanos y ajusta los rostros a una situación determinada. Durante las expresiones faciales, el robot utiliza el contacto ojo a ojo, expresión corporal y dialogo (velocidad, volumen, tono, etc.) para interactuar socialmente.

Los elementos principales de este método son la interrupción y la asignación previa de parámetros de movimiento (expresiones emocionales y ademanes); ambos permiten que el robot termine la interacción abruptamente para cambiar rápidamente de comportamiento según la situación. El sistema general se compone de seis elementos: 1) un organizador de datos estructurados que optimiza, almacena y relaciona la información que se procesa en forma de pares negativos-positivos (bueno-malo, feliz-triste, etc.), (“Key-Value Stores”) [106]; 2) un subsistema que decide la reacción del robot cuando sus sensores modifican la información que recibe (“Episode Rule Selector”); 3) un detector del número de usuarios parlantes (“Speaker Detector”); 4) un reconocedor de habla; 5) un planificador de movimientos y 6) un generador de diálogos. El comportamiento emocional en el robot no es el elemento central de la investigación; sin embargo, se hace un análisis extensivo en aspectos de interacción en contextos sociales reales.

Otro caso de estudio es el robot humanoide KOBIAN-R, el cuál utiliza 24 Grados de libertad para la activación de emociones; las expresiones son seleccionadas por un generador de rostros [15]. El sistema puede realizar más de 600,000 combinaciones a través de parámetros matemáticos

(clasificador polinomial); existen seis clasificadores para el generador de expresiones: 1) Modo; 2) Postura; 3) Temperamento; 4) Expectativa; 5) Certeza y 6) Poder de relación.

KOBIAN-R es capaz de activar expresiones en modo estático y dinámico; para aumentar o disminuir la intensidad de la emoción. Comúnmente el cambio entre expresiones faciales en los humanos es a través de transiciones de un estado emocional a otro con diferente intensidad y valencia. Una dinámico difiere de una estático en qué consiste en una secuencia de expresiones con intensidad incremental o en decremento [107, 108].

### **3.3. Pruebas de identificación facial**

Los prototipos faciales actuales integran modelos artificiales de emoción para generar expresiones; la gran mayoría están basados en las emociones básicas de Ekman y el sistema de codificación para la activación facial (“Facial Action Coding System” FACS.) [46, 109]. El sistema funciona como una herramienta y guía de clasificación para los movimientos faciales observables a través del análisis anatómico; descompone las expresiones faciales en componentes individuales del movimiento muscular. En esta guía se llaman Unidades de Acción (Action units, AU) a las áreas del rostro que se involucran al momento de generar gesticulaciones en lugar de nombrar a los músculos de la cara. “FACS” es altamente técnico; su manual describe el criterio de observación y codificación de cada unidad de acción (AU), además se describen las combinaciones posibles para cada unidad.

Algunos sistemas faciales se evalúan bajo un estricto método experimental; el más común es el de identificación de la expresión facial (“Facial Expression Identification”. FEI). En esta prueba el sujeto observa el rostro del robot (digital y/o físico) mientras ejecuta de manera aleatoria una serie predefinida de emociones (básicas universales); si es el caso, se presentan imágenes digitales sobre la pantalla. Durante la evaluación se hace una encuesta de tres reactivos; en el primero se le pide al sujeto que identifique la expresión; en el siguiente que evalúe su intensidad y en el último se da la oportunidad de identificar por segunda ocasión alguna expresión; el sujeto puede rectificar la respuesta del primer reactivo en una segunda opción. En el caso de que la prueba sea con un prototipo electromecánico, esta se inicia mostrando al robot con una cara neutral (ausencia de emoción), se activa el gesto emocional por algunos segundos y regresa al gesto neutral para alistar el siguiente cambio. El sujeto se limita a observar al rostro y en ningún momento existe algún tipo de interacción.

### **3.4. Enfoque minimalista**

El estudio de rostros robóticos se facilita con el desarrollo de diseños minimalistas reduciendo el equipo de construcción y aumentando la efectividad en la expresión de gestos. El prototipo facial MiRAE [14] es uno de los desarrollos cuyo diseño representa el enfoque minimalista con aplicaciones de alta eficiencia. El sistema ha sido probado con el instrumento de medición “FEI” para cada una de sus expresiones con resultados de legibilidad

muy altos. Las pruebas sugieren que los elementos como párpados, orejas, pómulos e incluso elementos que asimilan animales, no son requisitos para expresar gestos con alta legibilidad; específicamente los universales.

El sistema MiRAE tiene diez grados de libertad los cuales se consideran pocos para la ejecución de expresiones (dos para los ojos, dos para las cejas, cuatro para la boca y dos para el cuello). Los materiales de construcción son comerciales y algunos componentes son fabricados con impresión tridimensional. Su programación y control es a través de la placa electrónica Arduino (Genuino, <https://www.arduino.cc/>) que es un microcontrolador utilizado para facilitar el uso de la electrónica y sistemas; sus componentes hardware y software son de código abierto.

De acuerdo con el diseño de MiRAE, la posición de ojos, cejas y boca son suficientes para generar expresiones faciales. Previo a su construcción, se configuraron digitalmente los rostros emocionales y se realizaron pruebas de identificación facial; el resultado fue mejor en el diseño digital pues se consideró que es más fácil acoplar este tipo de gestos al sistema de codificación facial (FACS) [109]. Sin embargo, el movimiento de cuello en el prototipo físico incrementó la legibilidad del gesto; se obtuvieron resultados similares a su versión digital. La intensidad de la expresión se definió por la posición de los ángulos en los elementos del rostro; por ejemplo, las cejas colocadas al 50 % del total de grados que tiene el gesto con la intensidad más alta.

La ejecución de las expresiones digitales del robot MiRAE se presentan en la figura 3.3. En la figura 3.4 se muestran las expresiones sobre el

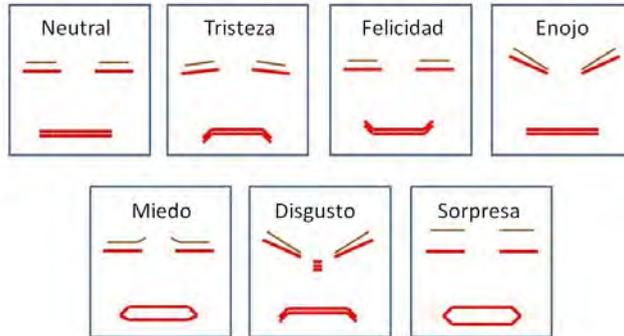


Figura 3.3: Expresiones faciales universales en MiRAE versión digital [14].

prototipo electromecánico (excepto el de disgusto).

Las pruebas realizadas respecto a la intensidad emocional tuvieron mejor resultado en la versión física; las expresiones con intensidad baja se hacen más legibles en estado dinámico. Los resultados obtenidos en el robot MiRAE se presentan en la figura 3.5; se comparan con los registros de los robots Kismet, BERT y el androide japonés Geminoid-F.

En general, el prototipo facial MiRAE se caracteriza por: a) Capacidad de expresión facial con pocos elementos (seis líneas horizontales), b) Reducción en la complejidad de construcción, c) Accesibilidad para el control de los rostros y d) Legibilidad en la ejecución de emociones básicas universales.

Las características del robot minimalista favorecen a su fabricación industrial. Esto los hace susceptibles a su desarrollo e implementación en robots interactivos con altas posibilidades para realizar evaluaciones controladas en poco tiempo.

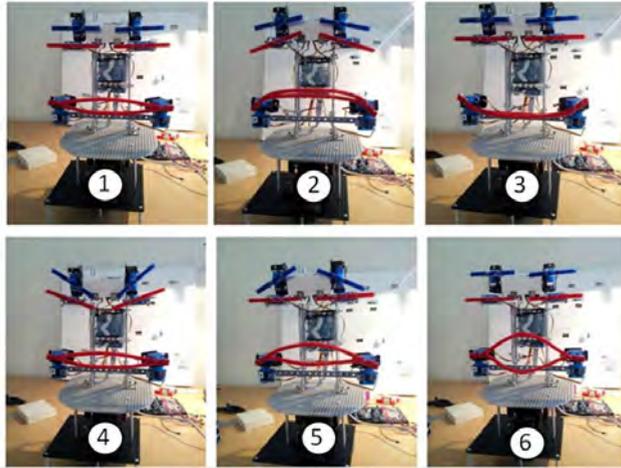


Figura 3.4: Expresiones faciales en MiRAE versión física [14]. 1.-Neutral, 2.-Tristeza, 3.-Felicidad, 4.-Enojo, 5.-Miedo y 6.-Sorpresa.

Expresión	MiRAE (n=30)	Kismet (n=17)	BERT (n=10)	Geminoid (n=71)
Felicidad	97%	82%	99%	88%
Tristeza	100%	82%	100%	80%
Enojo	87%	76%	64%	58%
Miedo	43%	47%	44%	9%
Sorpresa	97%	82%	93%	55%
<b>Promedio</b>	<b>85%</b>	<b>74%</b>	<b>80%</b>	<b>58%</b>

Figura 3.5: Promedio de identificación de la expresión facial (FEI) de las emociones de Ekman (excepto disgusto) para el robot MiRAE [14], Kismet [9], BERT [93] y Geminoid-F [95]. n= numero de sujetos.

### 3.5. Sumario

La implementación de rostros electromecánicos y digitales en robots interactivos permite analizar el impacto de las expresiones faciales en las personas además facilitan la interacción; siempre y cuando el diseño sea de acuerdo a las expectativas de los usuarios. La teoría de “Uncanny Valley” ofrece una referencia de criterios formales y constructivos.

Finalmente, a partir de las características generales de los prototipos faciales se han tomado en cuenta algunos aspectos que favorecen a su desarrollo y pruebas de funcionamiento correspondientes. A continuación, se enumeran estos criterios.

1. Fabricación de prototipos con enfoque minimalista; se reducen materiales y equipo electromecánico para su construcción.
2. Diseño abstracto para aumentar la legibilidad en la expresión sin comprometer la familiaridad y expectativas del usuario.
3. Uso de sistemas con hardware de código abierto (hardware libre) de fácil manipulación e implementación; susceptible a su activación a través de arquitecturas con modelos artificiales de emoción.
4. Versatilidad en el diseño; prototipo de fácil mantenimiento, con acceso rápido a los componentes para usarlo como un instrumento de medición y experimentación.
5. Sistema capaz de ejecutar expresiones faciales básicas.

6. Posibilidad de representar emociones con intensidades altas o bajas en formatos estáticos y dinámicos (transicionales).
7. Implementación del prototipo para pruebas de identificación facial.
8. Diseño susceptible para su desarrollo digital o físico con el fin de realizar experimentos en línea y presenciales.
9. Desarrollo de prototipos con estructuras robustas para integrarlos en robots interactivos dentro de entornos controlados y reales.
10. Diseño apto para su actualización.

## Capítulo 4

# Robot de servicio

La expresión de emociones ha sido implementada en robots utilizados para interactuar con personas o bien para realizar tareas específicas. En ambos casos la función primaria de los equipos es interactuar socialmente considerando el entorno (Diseño de espacios), los dispositivos implementados (Hardware) y capacidades propias del sistema (Arquitectura y software) [10, 110, 111].

La expectativa positiva hacía la robótica se deriva de su valor funcional; específicamente para servir a las personas. Con este fin, los robots se implementan dentro de plantas industriales en líneas de ensamble, armadoras, en áreas de transporte y carga, etc. Los sistemas que desempeñan tareas orientadas en entornos sociales hacen uso de elementos para interactuar directamente con personas; dialogo, seguimiento ocular, señalar, gestos corporales y faciales. Tales características facilitan la interacción y uso del sistema; un robot capaz de comportarse socialmente es visto con natura-



Figura 4.1: El robot enfermero Pearl [112, 113]. Interactuando en un asilo para ancianos.

lidad en espacios compartidos con las personas. Fong [10] define tres roles para los robots interactivos: Robots para asistencia a largo plazo, cómo colaborador y de interacción breve en espacios públicos.

## 4.1. Robots de asistencia

Los robots para asistencia son utilizados en espacios cómo hogares y áreas de atención médica. El robot enfermero Pearl [112, 113] (Figura 4.1), fue uno de los primeros sistemas probados en hospitales y asilos para ancianos; es capaz de navegar para funcionar cómo guía, reconocer voz y personas, emitir diálogos, además de contar con un comportamiento adaptativo; Pearl utiliza una arquitectura de tipo cognitivo.

## 4.2. Robots colaboradores

Otro rol de robots interactivos es como colaboradores; básicamente trabajando con humanos para alcanzar un objetivo en común. En este sentido el robot no es solo un asistente, sino que participa en la tarea de forma equivalente al colaborador; el sistema puede aprovechar las habilidades humanas cuando su software o hardware es limitado [114]. Cómo tal, el robot colaborador no se considera un par del humano, simplemente tiene una función específica y parcial dentro de la tarea. El sistema debe ser auto-suficiente en el nivel más básico y evitar que su acción le provoque algún daño; sin embargo, se espera que sea capaz de pedir asistencia al humano o comunicar con alguna señal social acerca de sus limitantes. En una configuración avanzada se espera que el robot tenga una conducta adaptativa.

## 4.3. Robots de interacción a corto plazo

Los robots de interacción a corto plazo en espacios públicos contrastan de los de asistencia y los colaborativos en que suelen actuar con humanos inexpertos; debido a esta desinformación, los usuarios tienen expectativas y conocimientos tecnológicos distintos. Otra diferencia sustancial radica en el tipo de interacción que se presenta en los espacios públicos la cual no está restringida a la relación uno a uno (humano y robot); el sistema no solo debe ser capaz de interactuar con varias personas a la vez, sino que también mantener la comunicación hasta satisfacer la necesidad del usuario [10]. Los sistemas robóticos en espacios públicos se diseñan como

guías en museos, para realizar actividades demostrativas en exposiciones, fuentes de información en estaciones de tren, distribuir panfletos en espacios abiertos entre otras.

Desde hace más de dos décadas se ha investigado acerca del manejo y control de la interacción con humanos a partir de la eficiencia del software y administración del hardware. Los estudios abordan los comportamientos en el robot tales como: manejo de dialogo, navegación, seguimiento y reconocimiento de personas, proximidad y distancias sociales, interacción múltiple (robot y personas), seguimiento ocular, expresión corporal, manipulación de objetos y expresiones faciales. Los desarrollos robóticos que han implementado este tipo de comportamientos son: El robot Rhino [115], Minerva [116], la serie de sistemas Mobot [117], Vikia [118], Robox [119], Robovie [120–122], Alpha [123](Figura 4.2), Robothino [124], Asimo [125], REEM [126, 127], Frog [128, 129] y Robovie R3 [130, 131] (Figura 4.3).

En la actualidad los sistemas integran capacidades de los tres tipos de los robots interactivos descritos. Esta combinación de características (asistencia, colaboración, coordinación, reacción, adaptación y relación social) se representan en dos clases bien identificadas en agentes artificiales: Robots sociales y robots de servicio.

#### **4.4. Robot social**

Generalmente un robot social se puede definir como un sistema físico, computacional y electromecánico, integrado al entorno social (dinámico y



Figura 4.2: Robots de interacción en espacios públicos. 1.- Rhino [115]. 2.- Mobot-Chips [117]. 3.- Mobot [117]. 4.- Vikia [118]. 5.- Robox [119]. 6.- Minerva [116]. 7.- Alpha [123].



Figura 4.3: Robots de interacción en espacios públicos. 1.- Robothino [124]  
 2.- Robovie [120–122] 3.- Frog [128, 129] 4.- REEM [126, 127] 5.- Robovie  
 R3 [130, 131] 6.- Asimo [125]

complejo); tiene un comportamiento reactivo dirigido al cumplimiento de objetivos definidos y al de los usuarios. Está capacitado para comunicarse socialmente con humanos y otros sistemas similares [132, 133]. El objetivo principal de su desarrollo es generar un comportamiento social inteligente utilizando señales naturales humanas [134]. Los proyectos de robótica social incluyen análisis científicos, técnicos y psicológicos. Los investigadores adquieren nuevos conocimientos a partir de la simulación del comportamiento humano en sistemas artificiales [135, 136] y creando interfaces que sean entendibles e intuitivas para el usuario [137]. Algunas investigaciones analizan los efectos de los robots sociales en humanos de forma psicológica y fisiológica con propósitos terapéuticos [138, 139]. Se sugiere que la observación y análisis del comportamiento durante la interacción humano y robot deberá estar situada en contextos sociales reales (“robots in the wild”); la interacción de robots en el medio social permite estudios objetivos y analíticos, con científicos entrenados y usuarios inexpertos [140]. El diseño de sistemas artificiales humanoides (Antropomorfismo) genera atributos favorables a la interacción; a nivel neurofisiológico y de comportamiento, las personas perciben a los robots con actitud e inteligencia similar al humano [141].

## 4.5. Robot de servicio

El concepto de robot de servicio se refiere a la entidad que es capaz de realizar un conjunto de comportamientos básicos utilizados durante la ejecución de tareas complejas [20]; es capaz de llevar a cabo actividades

en ambientes reales [142]. La estructura de su funcionamiento es gracias a su base de conocimientos o habilidades (comportamientos) los cuáles pueden representarse como un conjunto de procesos organizados (redes semánticas) [143,144]. Al igual que los sociales, los robots de servicio pueden interactuar con personas o llevar a cabo una colaboración social [145].

Los robots de servicio utilizan sensores y actuadores que les permiten percibir las características del entorno, visualizar objetos, personas y captar audio o voz, pueden navegar en el entorno, localizar áreas y manipular objetos. Los comportamientos de estos sistemas pueden ser de nivel básico dependientes de los datos directos de los sensores o complejos cuando el robot lleva a cabo tareas compuestas; un conjunto de comportamientos de nivel bajo que complementan un objetivo mayor (ej. “Trae una taza”). Para completar una tarea compleja el sistema estructura la actividad en una secuencia que le permita avanzar progresivamente hasta completar el objetivo [146]. Es deseable que un robot de servicio sea capaz de complementar una instrucción incompleta a través de representaciones y relaciones existentes en su base de conocimientos [147]. La integración de datos va desde los básicos operativos (bajo nivel) hasta la información útil para percibir y ejecutar sobre el entorno de interacción (alto nivel); el sistema ejecuta las acciones de acuerdo a su interpretación del mundo real [147].

El concepto y estructura funcional de un robot de servicio se determina a partir de la integración de elementos físicos que lo conforman (hardware) y el conjunto organizado de datos que corresponden a la programación, algoritmos y arquitecturas (software). El diseño del robot está definido por la

relación existente entre las entradas y salidas del sistema en la cual existe la transferencia de la información entre lo registrado y la respuesta correspondiente. El sistema de un robot de servicio se configura por: a) La capacidad de interpretar y reconocer representaciones externas en el entorno y b) Especificar y generar una respuesta a través de procesos cognitivos artificiales organizados por liberadores de acciones y comportamientos en relación al estado de la tarea y situación en el entorno. Con el fin de organizar sus comportamientos, los robots de servicio utilizan un lenguaje declarativo para establecer la estructura de la tarea y delimitar las acciones a realizar [20]. Los robots de servicio se encuentran ligados a la ejecución de tareas específicas; el sistema puede adoptarse ante situaciones y ofrecer respuestas de acuerdo a los estados en el entorno. Sin embargo, un sistema robótico expuesto a un ambiente variable reduce altamente su desempeño. Esto da pauta al desarrollo de robots inmersos en áreas de aplicación únicas; el estudio y observación del entorno aumentan la capacidad de ejecución del equipo y sistemas involucrados.

La idea general de la robótica de servicio doméstico es la de crear sistemas autónomos, multipropósito, capaces de interactuar a través de lenguaje natural y señales sociales no verbales. Con el fin de establecer el desarrollo organizado de sistemas autónomos se creó la liga “RoboCup@home” dentro de la competencia mundial “RoboCup” celebrada anualmente. En esta liga se pretende desplegar, ejecutar y probar la tecnología electro-mecánica/computacional para futuras aplicaciones domésticas y personales en entornos dinámicos no estandarizados presentes en el mundo real [148].

Durante la competencia existe la intención de identificar problemas en la implementación de sistemas y establecer puntos de referencia desde el criterio científico y tecnológico; básicamente se trata de acelerar y enfocar el concepto e integración de los robots de servicio en el ámbito cotidiano. A pesar de que la evaluación durante las pruebas es estricta, el diseño del robot y uso de la tecnología es altamente variable; el sistema involucra el software necesario para equilibrar su arquitectura y mantener estables sus capacidades (impacto en el rendimiento pero no es su capacidad) [20]. En la figura 4.4 se muestran algunos ejemplos de robots participantes en la liga “RoboCup@home”.

## 4.6. El robot Golem-III

El desarrollo del robot humanoide de servicio Golem-III (Figura 4.4, número ocho) forma parte de un proyecto nacional generado en el Instituto de Investigación en Matemáticas Aplicadas y Sistemas (IIMAS) de la UNAM. El proyecto inició desde 1998 por el Grupo Golem y se considera la integración de estructuras móviles con distintas capacidades; en la actualidad el sistema utiliza equipos que le permiten el registro y emisión de audio, manipulación de objetos, navegación, reconocimiento facial y corporal, captura de imágenes y vídeo. El principal objetivo de este proyecto es el de hacer una aportación a los elementos clave de la conversación humana, así como el desarrollo de esquemas de procesamiento complejo de datos basados en la generación de modelos de diálogo computacional



Figura 4.4: Robots de servicio “RoboCup@home”, “RoboCup” [148]. 1.- Cosero, Nimbro@home [149] 2.- Amigo, Tech United Eindhoven [150] 3.- Justina, Pumas@Home [151] 4.- Lisa, homer@UniKoblenz [152] 5.- Toyota HSR, Hibikino-Musashi@Home [153] 6.- Pepper, UChile Peppers [154] 7.- Golem-II+, The Golem Team [155] 8.- Golem-III, The Golem Team [157]

y la Arquitectura Cognitiva Orientada a la Interacción (IOCA) que permiten la rápida integración de sistemas con múltiples capacidades (visión, escucha, movimiento, síntesis de voz, etc.) [21,22,155–158]. Básicamente la interacción del sistema con el usuario y el medio ambiente guían el comportamiento del robot.

El uso de modelos de diálogo [159] para la integración de sistemas utilizados en un entorno son la base de ejecución de Golem-III; las tareas se completan a través de una serie de subsistemas o agentes del robot que cumplen con expectativas e intenciones del usuario y del propio sistema dentro de escenarios bien definidos. A través de modelos de diálogo se interpretan y simplifican las intenciones y acciones del robot; siguen un protocolo estructurado como secuencia de situaciones de conversación (ejecución del modelo de diálogo).

El uso de modelos de diálogo tiene varias ventajas [21]:

1. Los modelos se pueden representar mediante nodos y arcos; si una expectativa se cumple ( $e$ ), una acción será realizada ( $a$ ) y se pasará a la siguiente situación (ejemplo: Situación  $i$  a Situación  $j$ ) (figura 4.5).
2. Las acciones, expectativas y situaciones se pueden especificar por predicados que incluyan variables libres o funciones proposicionales.
3. Es posible asignar estados recursivos a las situaciones que forman parte de un modelo de diálogo completo (figura 4.6). La recursividad permite reiniciar el ciclo de interacción omitiendo las situaciones, acciones y expectativas que se han cumplido con el fin de completar

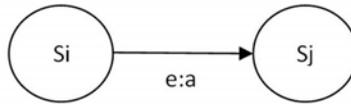


Figura 4.5: Representación gráfica de un modelo de diálogo. Se pasa de la situación  $i$  a la situación  $j$  si se cumple una expectativa ( $e$ ) y una acción es realizada ( $a$ ).

la tarea sin repetir completamente la rutina; reutiliza la información con parámetros diferentes.

4. Pueden asignarse estados de recuperación en caso de que la tarea sea difícil de llevarse a cabo y el robot pueda continuar interactuando con el usuario.
5. Permite el acceso a la historia de la interacción

Golem-III tiene varios comportamientos y habilidades (figura 4.7); pueden combinarse según el tipo de tarea a realizar. Un modelo de diálogo puede estar compuesto por situaciones embebidas dentro de un modelo superior; es decir jerarquías conceptuales para cumplir con alguna expectativa a través de sub-modelos de diálogo. La figura 4.8 muestra un ejemplo de jerarquías estructuradas en un modelo de diálogo en donde se pide al robot la instrucción “find” [21]. En la figura, el modelo para “find” (encontrar algo) hace uso del comportamiento “scan” (movimiento horizontal de la cámara del robot) y “tilt” (movimiento vertical de la cámara del robot), una vez realizado un número determinado de observaciones en el entorno y

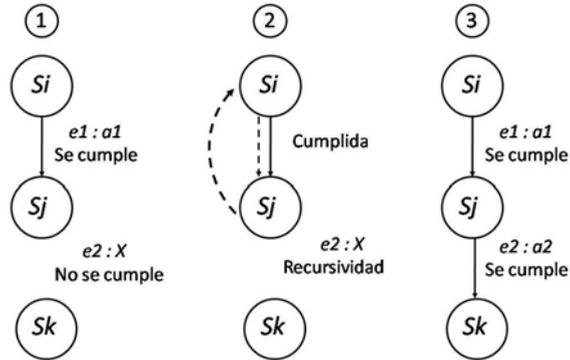


Figura 4.6: Recursividad en los modelos de diálogo. En el caso (1) la expectativa y acción se cumplen de  $S_i$  a  $S_j$ , no así de  $S_j$  a  $S_k$ . En el caso (2) se realiza la recursividad omitiendo repetir la situación cumplida. En el último caso se cumple la expectativa faltante y se completa la tarea.

ejecutando a su vez el comportamiento “see”, se selecciona un cierto tipo de elemento (“object”, “face”) y el tipo de observación deseada (“detect”, “recognize”). Este modelo de diálogo se organizan comportamientos básicos (estáticos) para completar una tarea compleja compuesta (dinámicos).

La Arquitectura Cognitiva Orientada a la Interacción (IOCA) [21, 22], interpretada a través de los modelos de diálogo, permite configurar la interacción entre el robot y el entorno. El ciclo de interacción involucra los siguientes módulos: a) Reconocimiento, b) Interpretación, c) Modelos de diálogo, d) Especificación y e) Renderización. La arquitectura permite ejecutar un modo reactivo en el robot en caso de que algún fenómeno no previsto en el entorno (ni en la secuencia de tarea) interfiera con la ejecución del modelo de diálogo. El sistema reactivo utiliza un módulo denomina-

approach object/person/surface	find	recognize: body/face
ask	grasp	relieve
consult	guide	say
deliver	memorize: body/face	scan
detect: door/face/head	move	see: face/gesture/object/person
doas	plane_detect	take
execute	prompt	tilt

Figura 4.7: Comportamientos y habilidades del robot Golem-III. Con los comportamientos individuales es posible generar comportamientos compuestos

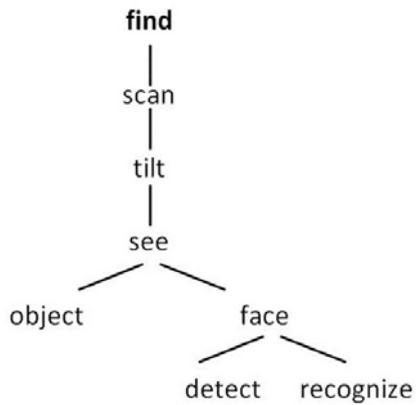


Figura 4.8: Modelo de diálogo compuesto para la conducta “find” y comportamientos individuales.

do “coordinador” para administrar su ejecución con la de los Modelos de Diálogo sin interrumpir ambas funciones. En la figura 4.9 se muestra la Arquitectura IOCA coordinada con el Sistema Reactivo Autónomo.

Gracias a los modelos de diálogo, el robot Golem-III es capaz de realizar varias tareas de servicio; se espera que las expresiones emocionales a través de un prototipo facial electromecánico mejoren el sistema y la interacción humano-robot en los siguientes aspectos:

1. Las expresiones emocionales podrán activarse como un comportamiento más dentro de los modelos de diálogo del sistema.
2. Informarán el estado interno del sistema.
3. Servirán para comunicar al usuario un estado en la tarea.
4. Se establecerá una relación entre el emoción e historia de la interacción a través de la intensidad del gesto y su estado dinámico (transitivo).
5. Se pretende que la expresión influya en la atención y las acciones del usuario de forma positiva a la interacción, así como a favor de la tarea.

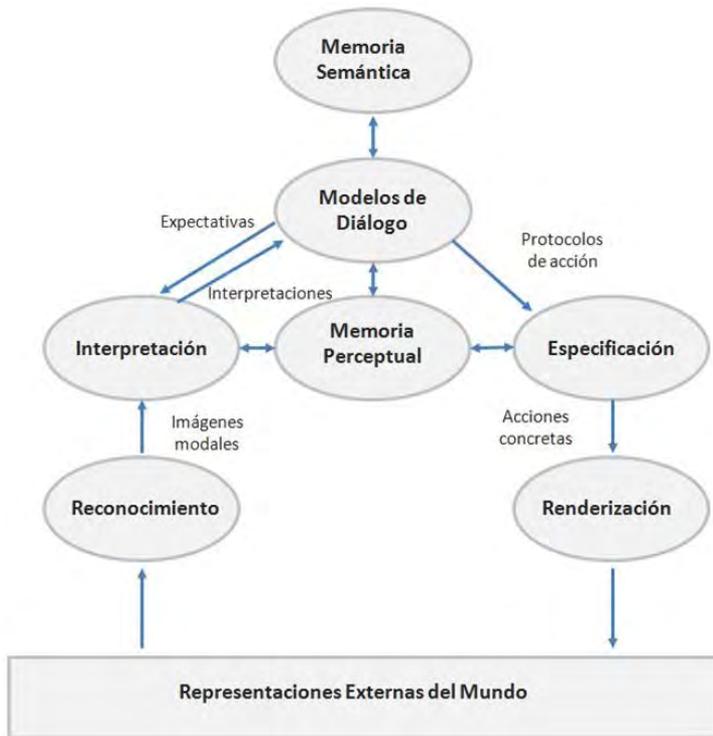


Figura 4.9: Representación gráfica de la Arquitectura Cognitiva Orientada a la Interacción (IOCA) [21, 22].



## Capítulo 5

# Prototipo facial y pruebas de identificación

El robot Golem-III es capaz de realizar varias tareas de servicio y de interactuar con humanos por medio del dialogo, reconocimiento de rostros, seguimiento de personas y manipulación de objetos. La integración de un prototipo facial en el robot contribuirá a la interacción y comunicación con el usuario; en particular se analizará la influencia de las expresiones faciales sobre el usuario dentro del contexto de tareas colaborativas.

El estudio de las expresiones faciales durante las tareas del robot se realizó a partir de dos prototipos faciales; uno de ellos para experimentación y el otro para implementarlo en el robot Golem-III. El diseño de ambos prototipos faciales se realizó bajo los criterios especificados al final del capítulo tres, sin embargo, para el prototipo que fue implementado en

Golem-III se tomó en cuenta el diseño de imagen y configuración formal para favorecer la interacción. Los detalles de cada prototipo fueron diferentes en el aspecto formal; por un lado, el diseño minimalista en la versión experimental y por el otro, el diseño con elementos de mayor detalle para generar expectativas positivas en el usuario. El sistema facial es capaz de activarse a partir de los dispositivos de hardware y software en el robot Golem-III.

## 5.1. Prototipo facial GolemX-1

El rostro robotico GolemX-1 se diseñó como prototipo de pruebas; es un sistema facial minimalista capaz de ejecutar las emociones descritas por Ekman [3,17]. En su diseño se tomó en cuenta el mínimo de elementos para la expresión de gestos [14] y el uso de ocho grados de libertad (“Degrees of Freedom” DOF). Se fabricó el prototipo con equipo electromecánico y hardware de código abierto (hardware/software libre) de fácil manipulación, implementación y de bajo costo [160,161].

La estructura principal del sistema esta basada en el principio de fabricación por prototipaje rápido (“Rapid Prototyping”); proceso para crear estructuras físicas tridimensionales a partir de sistemas de computación (CAD/CAM) en tiempo corto utilizando sistemas para dibujo y maquinado de piezas en diversos materiales [162–165]. La estructura principal del prototipo facial se realizó con plástico acrílico cortado a láser [166,167], con piezas para detalles y ensamblajes en impresión tridimensional (depo-

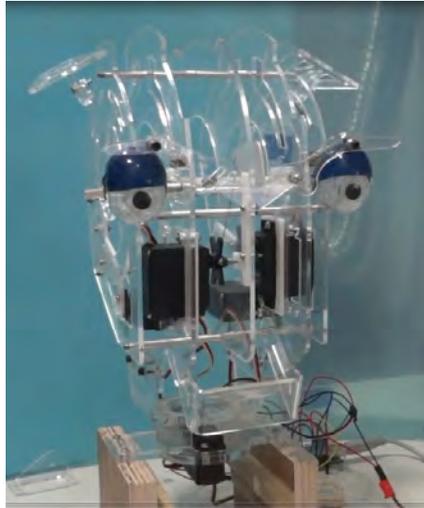


Figura 5.1: GolemX-1. Prototipo versión *Beta*.

sición de material) con el polímero poliácido láctico (PLA) [168]. Ambos son un proceso de bajo costo para elaborar piezas replicables y de fácil mantenimiento.

En la figura 5.1 se muestra la primera versión del prototipo de pruebas el cual incluye el siguiente equipo electromecánico: a) Una placa Arduino UNO, b) Cinco servo motores marca Futaba modelo *S148* y c) Una batería Futaba *NR-4J*.

El prototipo final de GolemX-1 fue simplificado en el aspecto constructivo y de manufactura; para tal efecto se diseñó previamente una versión digital la cuál fue sujeta a evaluación facial en un experimento en línea cómo se muestra más adelante [49]. Se modelaron digitalmente las siguientes expresiones: Neutral, felicidad, enojo, tristeza, sorpresa, miedo y preocupado

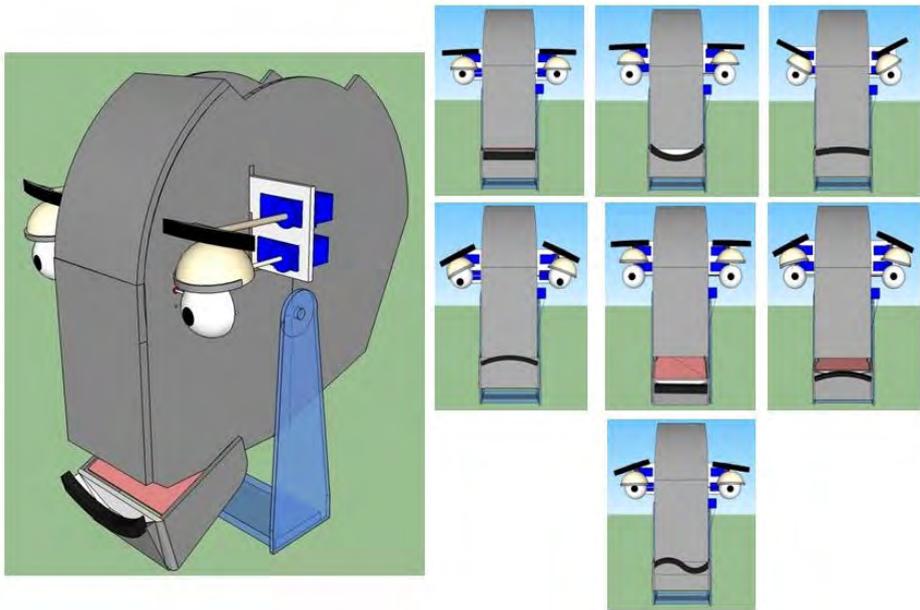


Figura 5.2: Diseño digital de GolemX-1 (izquierda). Expresiones emocionales (derecha). De izquierda a derecha y arriba hacia abajo: Neutral, felicidad, enojo, tristeza, sorpresa, miedo y preocupado

(Figura 5.2).

El diseño digital facilitó la fabricación del prototipo físico; en comparación con la versión *Beta*, se redujo el material para su fabricación, se simplificó el proceso de armado, la representación de las emociones fue mejor (según encuestas realizadas entre colegas) y se facilitó la implementación del equipo electromecánico. Sin embargo, se utilizaron siete grados de libertad (7DOF) para la ejecución de gestos; dos más que en la versión anterior. Se utilizó el mismo proceso de manufactura y materiales. En la figura 5.3

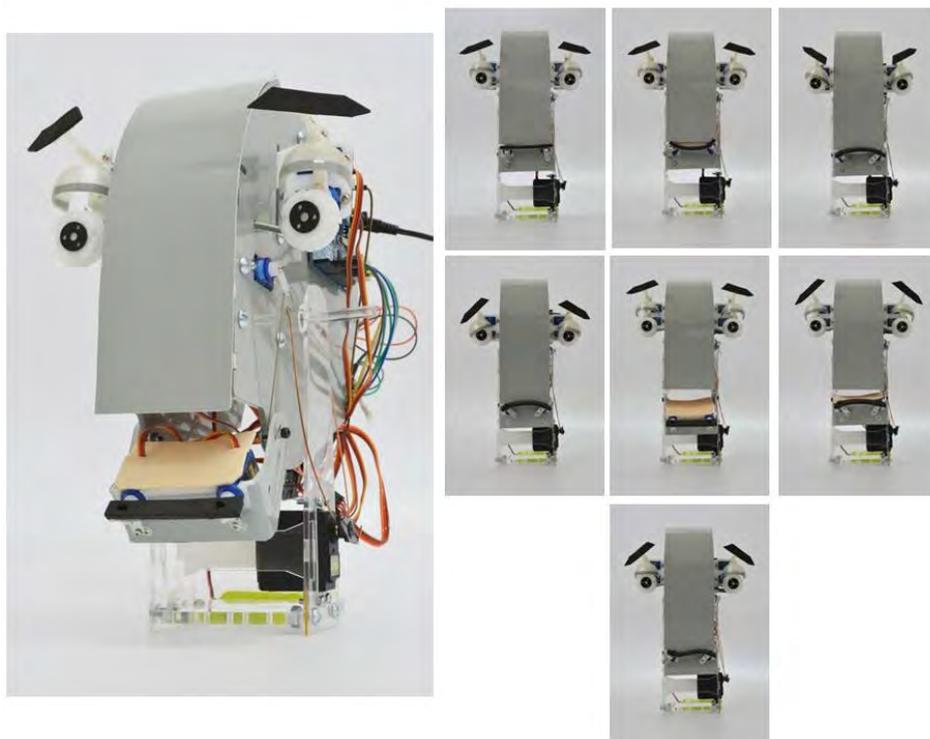


Figura 5.3: Prototipo físico de GolemX-1 (izquierda). Expresiones emocionales (derecha). De izquierda a derecha y arriba hacia abajo: Neutral, felicidad, enojo, tristeza, sorpresa, miedo y preocupado

se muestra la versión final del prototipo el cual incluye el siguiente equipo electromecánico: a) Una placa Arduino UNO, b) Cinco servomotores marca TowerPro *Sg90* de 9 gramos, c) Dos servomotores marca TowerPro *Mg995* y d) Una batería Futaba *NR-4J*.

## 5.2. Pruebas de identificación facial

EL prototipo facial fue sujeto a pruebas de legibilidad tanto en su versión digital y física; ambas con un espacio de dos años entre sus evaluaciones. Se realizó la evaluación de legibilidad de expresiones con las imágenes en la figura 5.2; a partir de los resultados se desarrolló el prototipo físico y se realizaron las pruebas correspondientes.

Las evaluaciones de legibilidad en ambas versiones (digital y física) se realizaron a partir de un estudio en línea con opción múltiple; se presentó el rostro neutral, de felicidad, enojo, tristeza, sorpresa, miedo y preocupado a estudiantes de diseño industrial y de ingeniería en computación en sesiones individuales. El promedio de edad fue de 24 años. Los sujetos no habían visto el prototipo con anterioridad. Los siete gestos se presentaron en pantalla; uno a la vez. El sujeto seleccionaba el nombre de la expresión de una lista con siete etiquetas emocionales (Se le pidió al sujeto marcar la etiqueta apropiada). Las etiquetas se presentaron en la pantalla a un lado de la imagen. Tanto las imágenes del rostro y etiquetas se presentaron aleatoriamente. La imagen cambiaba cada vez que el usuario indicaba el tipo de expresión. En la figura 5.4 se presenta la interfaz para la evaluación

en línea.

Para el estudio de identificación utilizando la versión digital, 88 sujetos respondieron a la invitación; en el caso del rostro físico, 98 personas diferentes llevaron a cabo el experimento. Los resultados de la prueba para la versión digital se presentan en el cuadro 5.1. La matriz de confusión muestra que las expresiones negativas tuvieron una precisión mayor que las positivas; la expresión de enojado obtuvo la legibilidad más alta (94,3%), mientras que el de tristeza fue el segundo mejor reconocido (85,2%), neutral el siguiente (76,1%) por delante de las expresiones positivas de felicidad (63,6%) y sorpresa (59,1%). La expresión de miedo obtuvo la calificación más baja (12,5%); en general es difícil de identificar según otras pruebas de legibilidad [14].

En el caso de la expresión de preocupado se registró una calificación relativamente baja (51,1%); posiblemente derivado del uso de los labios torcidos (la posición de la esquina de la boca en dirección contraria de la otra esquina similar a la letra “S” en sentido horizontal). El uso de los labios torcidos como expresión facial no está estandarizado para la expresión de preocupación y no se encuentra registrada en el sistema de codificación para la activación facial (FACS) [46].

El cuadro 5.2 resume los resultados de la encuesta para el prototipo físico; los datos son consistentes con los hallazgos anteriores: el enojo obtuvo la mayor precisión de identificación (95,9%), tristeza la segunda posición (93,8%), seguido de neutral (82,6%) y feliz (74,4%), que presentó un aumento, la sorpresa (59,1%) obtuvo el mismo resultado. El rostro

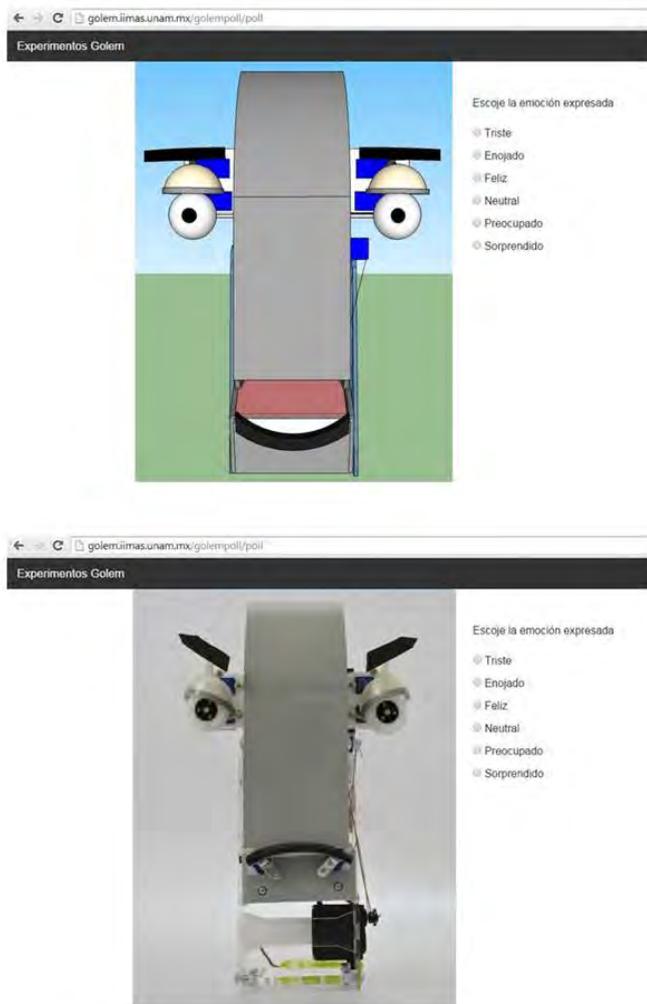


Figura 5.4: Interfaz para pruebas de legibilidad del prototipo GolemX-1, Imagen y etiquetas de opción múltiple. Formato digital (Arriba) y Formato físico (abajo)

Cuadro 5.1: Matriz de confusión para el rostro digital (88 sujetos). Enojo:**E**, Feliz:**F**, Miedo:**M**, Neutral:**N**, Triste:**T**, Sorpresa:**S** y Preocupado:**P**

Opciones	Imágenes							
	<b>E</b>	<b>F</b>	<b>M</b>	<b>N</b>	<b>T</b>	<b>S</b>	<b>P</b>	<b>Precisión</b>
<b>E</b>	<b>83</b>	3	3	11	1	4	8	94.3 %
<b>F</b>	1	<b>56</b>	0	0	1	4	0	63.6 %
<b>M</b>	0	1	<b>11</b>	0	1	3	10	12.5 %
<b>N</b>	2	11	2	<b>67</b>	3	17	14	76.1 %
<b>T</b>	1	0	32	6	<b>75</b>	2	9	85.2 %
<b>S</b>	0	16	20	1	1	<b>52</b>	2	59.1 %
<b>P</b>	1	1	20	3	6	6	<b>45</b>	51.1 %

preocupado fue mejor reconocido (58,1 %) y el miedo (11,2 %) sigue en la parte inferior de las puntuaciones. En general, los resultados mostraron un aumento de la legibilidad en comparación con su versión digital; esto se atribuye a los ajustes realizados sobre el prototipo electromecánico y al efecto físico tridimensional en las imágenes.

### 5.3. Prototipo facial de Golem-III

El diseño del prototipo facial de Golem-III estuvo sujeto al acoplamiento con los sistemas y equipo funcional del robot. De igual manera que el prototipo de pruebas, se utilizó equipo de control electrónico de software abierto, la ejecución de expresiones faciales universales y elementos básicos

Cuadro 5.2: Matriz de confusión para el prototipo físico (98 sujetos).

Opciones	Imágenes							Precisión
	E	F	M	N	T	S	P	
E	<b>94</b>	1	0	0	2	0	3	95.9%
F	0	<b>73</b>	4	4	0	16	1	74.4%
M	0	0	<b>11</b>	1	1	1	6	11.2%
N	1	13	5	<b>81</b>	1	17	9	82.6%
T	1	4	26	1	<b>92</b>	0	19	93.8%
S	0	5	28	5	1	<b>58</b>	3	59.1%
P	2	2	24	6	1	6	<b>57</b>	58.1%

para su manufactura. El sistema facial permite el funcionamiento de distintos dispositivos de identificación, visión y audio. Por otro lado, el peso del prototipo facial no excede los 500 gramos; se evita la sobrecarga de los actuadores responsables del movimiento de la estructura superior del robot (movimiento de cuello).

El diseño en los sistemas Golem mantiene algunas similitudes en cuanto a sus rasgos, colores, logotipos y figuras representativas de la UNAM; las imágenes de la figura 5.5 muestran el desarrollo e ideas generales del producto final deseado. Las primeras pruebas del rostro implementadas en Golem-III se muestran en la figura 5.6; en el diseño se evitó obstaculizar los dispositivos de funcionamiento (Sensor Kinect) lo que alteró el concepto y ejecución de gestos emocionales.

El prototipo final implementado en el robot se muestra en la figura 5.7.

El sistema cumple con los requisitos necesarios tanto físicos, constructivos (peso, materiales, procesos, etc.) y funcionales (activación a través de modelos de diálogo, software así como sistemas de código abierto y el uso de seis servomotores marca TowePro SG90). El mantenimiento y cambio de componentes son sencillos de realizar

La evaluación de los rostros se realizó con el mismo formato que en el prototipo de pruebas; se presentan los resultados en el cuadro 5.3. En esta ocasión 56 personas que no realizaron los experimentos anteriores llevaron a cabo el experimento. El promedio de edad fue de 20 años. Se observó una variación general en la legibilidad de las expresiones en comparación al prototipo de pruebas excepto en el de enojo que fue identificado por todos los sujetos (100%), sorpresa obtuvo la segunda posición (92,9%), seguido de felicidad (89,3%), después tristeza (76,8%), el cuál registró un descenso. En el caso de miedo y neutral obtuvieron la misma puntuación (62,5%). Cabe destacar el considerable aumento en la legibilidad de miedo en comparación a los resultados en las pruebas anteriores; hipotéticamente, las variantes en los resultados se atribuyen directamente a la configuración formal, sin embargo el rostro de enojo continua siendo el más legible. En esta ocasión se omitió la cara preocupada. El estudio se realizó para conocer el reconocimiento del rostro en un prototipo totalmente diferente al realizado en las pruebas anteriores; no es posible comparar directamente el resultado ya que el número de sujetos fue inferior, sin embargo se puede observar una tendencia favorable al desempeño general en la representación de emociones (excepto tristeza).



Figura 5.5: Concepto inicial del prototipo facial para Golem-III

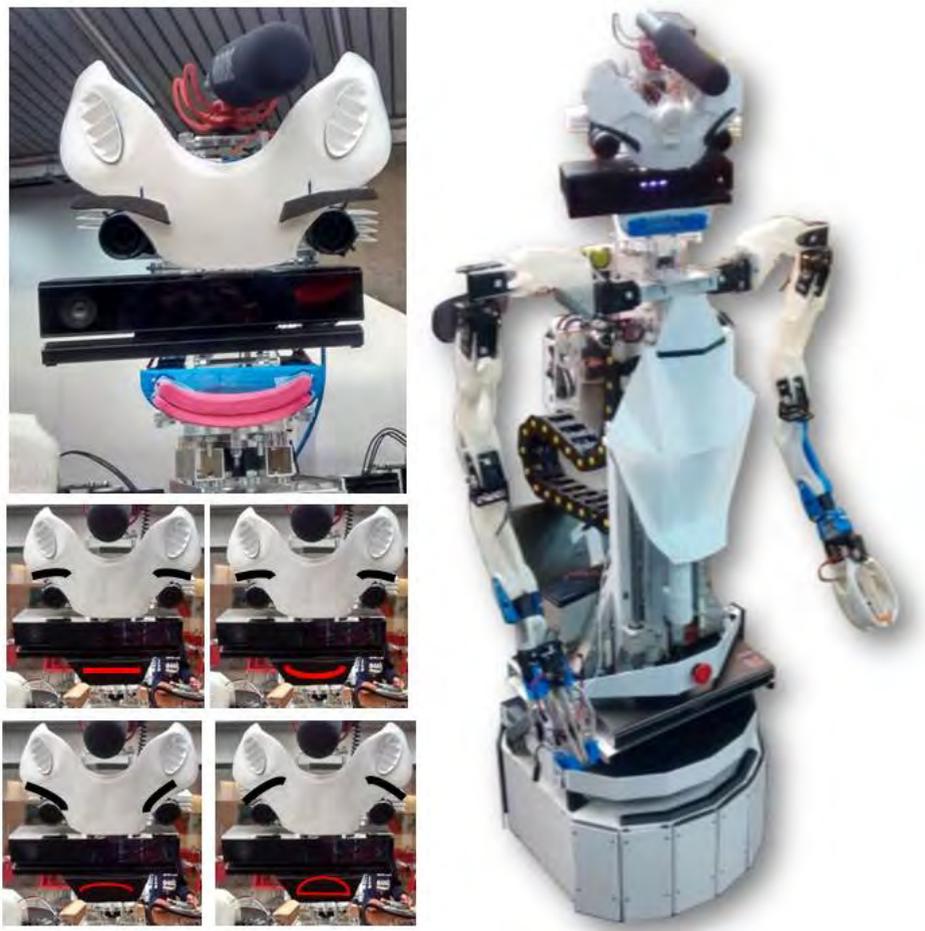


Figura 5.6: Prototipo facial *Beta* para Golem-III

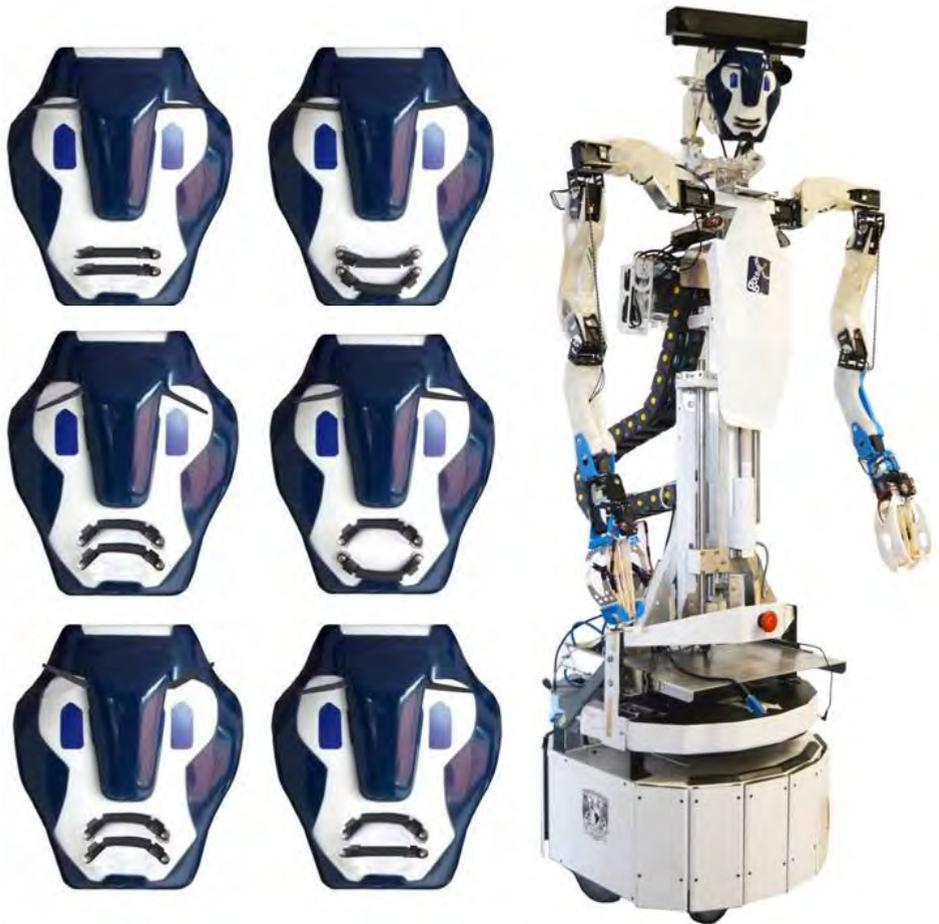


Figura 5.7: Prototipo facial final para Golem-III. Expresiones faciales: (Izquierda a derecha) Neutral, feliz, tristeza, sorpresa, enojo y miedo

Cuadro 5.3: Matriz de confusión para el prototipo facial de Golem-III (56 sujetos). Enojo:**E**, Feliz:**F**, Miedo:**M**, Neutral:**N**, Triste:**T** y Sorpresa:**S**

Opciones	Imágenes						Precisión
	<b>E</b>	<b>F</b>	<b>M</b>	<b>N</b>	<b>T</b>	<b>S</b>	
<b>E</b>	<b>56</b>	0	1	0	0	1	100 %
<b>F</b>	0	<b>50</b>	0	0	0	0	89.3 %
<b>M</b>	0	2	<b>35</b>	16	13	1	62.5 %
<b>N</b>	0	4	0	<b>35</b>	0	1	62.5 %
<b>T</b>	0	0	8	4	<b>43</b>	1	76.8 %
<b>S</b>	0	0	12	1	0	<b>52</b>	92.9 %

Las pruebas de expresión facial mostraron que tanto en el prototipo de pruebas digital, el electromecánico y el implementado en Golem-III, la expresión de enojo fue el más legible seguido del de tristeza. En el caso del prototipo final, fue el de felicidad el segundo mejor reconocido. Se considera que el concepto formal influyó directamente en la legibilidad a pesar de que la posición de los componentes (cejas y boca) fueron semejantes; para el de sorpresa se realizó un cambio en la posición de la boca. Los resultados tuvieron una precisión similar en las evaluaciones del prototipo de pruebas, sin embargo el final presentó una diferencia general en los resultados; sorpresa y miedo incrementaron su porcentaje. El cambio en la configuración formal de la boca en la expresión de sorpresa favoreció la legibilidad (del 59.1 % al 92.9 %). Sin embargo, en las tres evaluaciones de los cuadros 5.1, 5.2 y 5.3 presentan al rostro de miedo con menor precisión

en la identificación. En general, la identificación de la expresión de miedo en robots tiene un porcentaje bajo, tal cómo se muestra en la figura 3.4 en el capítulo 4, con la evaluación de expresiones faciales en otros sistemas. Particularmente, se considera que tal expresión se le atribuye incorrectamente debido a la falta de elementos que complementan la expresión, cómo por ejemplo el entorno; es probable que la emoción tenga mayor sentido si se muestra en lugares que podrían evocarla, cómo por ejemplo, un lugar oscuro, una situación que ponga en riesgo la integridad física, etcétera. Además, la mala legibilidad es independiente del tipo de diseño y método constructivo del rostro robótico.

## 5.4. Modelo de diálogo

El modelo de diálogo para la activación de las expresiones faciales se origina a partir de la interpretación de las situaciones externas al sistema, el procesamiento de la información y respuesta a través de la activación del hardware. En la figura 5.8 se presenta de manera gráfica la estrategia de implementación del modelo de emociones y generación de expresiones faciales en el robot.

Para generar el modelo de diálogo (dm) se creó un comportamiento denominado “*mood\_dm*”. La complejidad del modelo radica en la asignación de emociones a las situaciones exitosas, fallidas, no esperadas y cuando se requiera advertir u orientar.

Durante el diseño del modelo de diálogo se tomó en cuenta la orga-

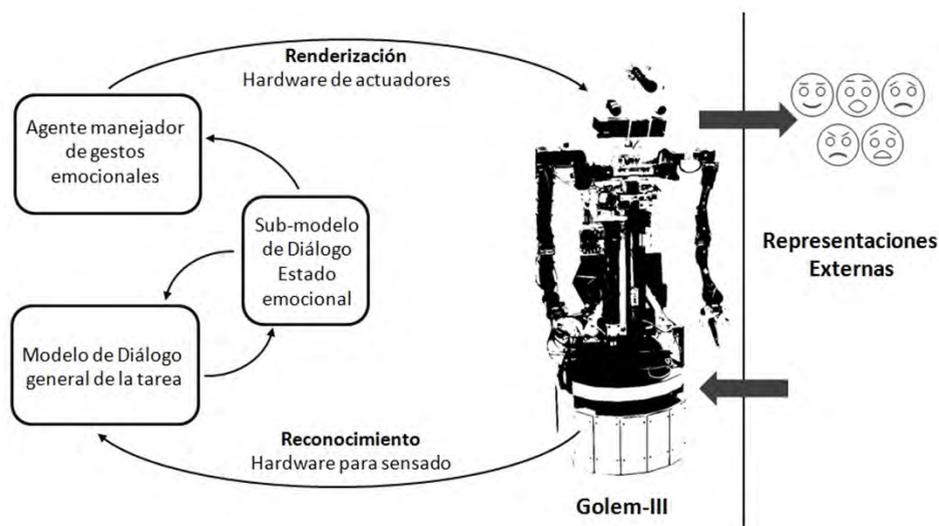


Figura 5.8: Implementación gráfica del modelo de diálogo en Golem-III.

nización estructural de tareas la cual está compuesta por expectativas y acciones que resultan en situaciones diferentes. Este método permite asignar el gesto según el efecto a generar en la tarea; el modelo de diálogo puede llamar a cualquier expresión que sea útil a la situación sin obstaculizar los modelos de diálogo que se incorporan durante la tarea. El modelo se representa gráficamente en la figura 5.9.

La figura 5.10 muestra el uso del modelo de diálogo “*mood\_dm*” dentro de una tarea simple. La representación gráfica corresponde al sub-modelo de diálogo en la tarea “Follow Me” que integra el comportamiento para la ejecución de las expresiones faciales. Al inicio de la actividad se sigue utilizando una cara neutral con el fin de que el cambio en el comportamiento emocional sea evidente.

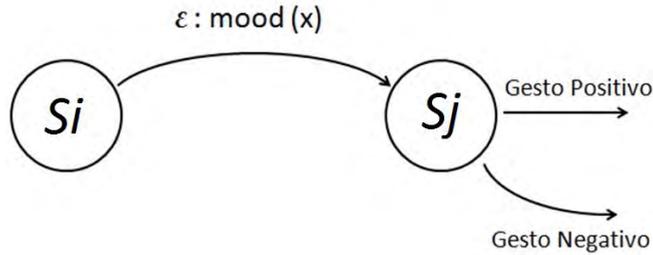


Figura 5.9: Representación gráfica del modelo de diálogo “*mood\_dm*” [21, 22]. El modelo genera a partir de una expectativa inicial y la acción “*mood*” ( $\varepsilon:mood$ ) para activar el gesto emocional; en este caso  $x$ , por ejemplo: *mood (neutral, happy, sad, angry, surprise, fear)* que corresponde al gesto solicitado para ser asignado al agente generador de gestos.

## 5.5. Sumario

Las tareas del Robot Golem-III y las de los sistemas predecesores se realizan en laboratorio, en demostraciones abiertas al público y dentro de la competencia RoboCup. En estos tres entornos se presentan algunas situaciones y acciones recurrentes; el robot puede ejecutar expresiones faciales para mostrar su estado interno y la postura emocional respecto a su representación del estado del externo.

Las acciones especificadas en la figura 5.11 pueden ser ejecutadas durante tareas de colaboración humano-robot. Las arquitecturas y modelos computacionales para la activación de emociones en robots se han evaluado e implementado durante la interacción con personas [169–172] sin embargo el efecto de las expresiones faciales sobre los usuarios se ha estudiado de

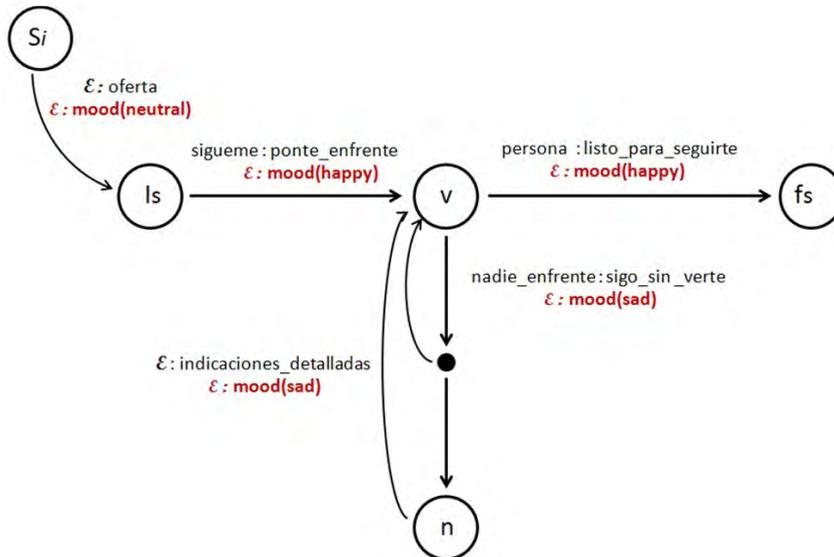


Figura 5.10: Representación gráfica de la tarea “Follow me”. El robot inicia la tarea (*Si*) emitiendo un mensaje indicando que es capaz de seguir a una persona; se ejecuta la cara neutral. En la situación (*Is*) el robot espera la indicación “sígueme” del usuario, una vez registrada, la siguiente acción del robot es pedir al usuario colocarse delante; se activa el gesto de felicidad. En la situación (*V*) el robot activa su interfaz de vídeo para identificar al sujeto, si la acción es fallida el sistema intenta repetir esta acción una vez más y hasta tres como máximo; por cada intento el sistema activa el rostro de tristeza. En caso de que el sistema no registre, el robot dará instrucciones específicas; durante la situación (*n*) se indica detalladamente a la persona el procedimiento para ser identificado o en su defecto se da por terminada la interacción. En caso de registro, el robot se prepara para seguir al usuario; la cara de felicidad continua o se incrementa la intensidad de la emoción. El robot sigue a la persona (situación *fs*).



Figura 5.11: Gestos emocionales robóticos y posibles escenarios para su activación durante tareas colaborativas.

forma moderada; sobre todo durante la tareas de agentes autónomos de servicio, sociales y de colaboración.

Las emociones activadas en los robots orientan al usuario acerca del estado interno del sistema; muchas de las emociones son asignadas de forma arbitraria y pretenden sacar provecho favor de la interacción. Con el fin de analizar el efecto de las emociones en las personas durante tareas de servicio y colaborativas, se diseñó un prototipo facial capaz de activar expresiones universales (excepto disgusto). De acuerdo a los experimentos de identificación facial los rostros negativos tienen mayor legibilidad; por tal razón son susceptible a su activación y estudio durante la interacción humano-robot.

Las expresiones emocionales pueden favorecer a la tarea del robot de

servicio siempre y cuando se pueda anticipar, conocer e incluso predeterminar la respuesta del usuario en relación a la situación, condición del entorno y tipo. La meta principal de esta investigación es la de establecer las bases para el uso de expresiones emocionales durante las tareas de un robot de servicio en colaboración con el usuario; se analizará el efecto positivo de los rostros de mayor legibilidad (negativos) en las acciones y atención humana.

Finalmente, la activación de expresiones emocionales en Golem-III involucra los siguientes aspectos:

1. Desarrollo de técnicas para representar expresiones transitorias entre las categorías discretas (ej. Triste a feliz, negativo a positivo).
2. Estudio de la relación entre expresiones faciales transitorias con las estáticas.
3. Análisis de la influencia de los rostros emocionales robóticos sobre el usuario. - Influencia en las acciones humanas. - Influencia en la atención humana.



## Capítulo 6

# El efecto positivo de una retroalimentación negativa

El prototipo facial GolemX-1 es legible en las expresiones negativas tristeza y enojo tal como se mostró en las pruebas de identificación facial; esta condición favoreció al análisis durante la interacción humano-robot y la confluencia con los gestos positivos (felicidad y sorpresa). Básicamente se realizó el análisis del efecto de las expresiones faciales en los humanos a favor de una tarea. Las actividades colaborativas entre humano y robot se implementaron durante un segundo experimento en donde se evaluaron las reacciones y comportamientos del sujeto como respuesta a la retroalimentación de emociones negativas en el robot (retroalimentación negativa).

El uso de las señales sociales en sistemas robóticos es crucial para el estudio de la interacción entre humano y robot. Estas señales son útiles para

coordinar sus acciones; sirve a la interpretación del razonamiento humano y de patrones cognitivos. Además, sirve para medir y cuantificar el impacto de la interacción en las personas [173]. La conducta, apariencia física, lenguaje natural, seguimiento ocular, señalar, gesticulación, personalidad y lenguaje corporal en los robots influyen directamente en la percepción y comportamiento humano. Las señales sociales no verbales en sistemas artificiales mejoran la interacción con humanos; por ejemplo, el seguimiento ocular en robots puede advertir e indicar precaución [28].

El uso de señales sociales ha sido estudiado ampliamente a través de rostros artificiales para coordinar y fortalecer la interacción así como predecir las intenciones entre colegas (humano y robot) [28, 174]; es posible influenciar la precisión y tiempo de respuesta en humanos [27]. Dentro del área manufacturera se ha demostrado que el uso de gestos en robots industriales mejora la comunicación con los usuarios. Las expresiones robóticas tienen influencia positiva durante tareas colaborativas industrializadas reflejadas en la continuidad de las actividades, reduciendo el periodo de trabajo e incremento de la productividad [175].

En general, las expresiones faciales negativas y positivas se perciben de diversas maneras por los humanos; la identificación de tristeza tiene mayor influencia en la atención humana que el de felicidad, sin importar que la identificación sea intencionada o indirecta [26, 56]. La retroalimentación negativa en agentes artificiales ha resultado eficaz durante tareas con humanos [57].

## 6.1. El efecto regulador de las expresiones faciales negativas

La influencia de la tristeza y el enojo se analizaron en un experimento de interacción humano-robot; se cuestionó si la presencia de una emoción negativa contribuye a completar correctamente una tarea. Para el análisis del efecto de rostros negativos se diseñó una interacción colaborativa; el humano y robot llevaron a cabo una tarea repetitiva no compleja en la cual cooperaban para colocar diez objetos cilíndricos dentro de un contenedor (el humano pasaba el objeto al robot para lanzarlo al contenedor). El robot estaba programado para interrumpir repentina y aleatoriamente su colaboración; en este caso, el sistema no era capaz de lanzar el objeto al contenedor. Durante la interrupción, las expresiones negativas se activaron para anunciar la incapacidad del sistema; en una sesión se presentó tristeza y en otra el de enojo. Para contrastar la valencia en la activación del gesto (positivo a negativo), la expresión de felicidad se mostró cuando el robot lograba colocar el objeto dentro del contenedor. El comportamiento emocional en el sistema se comparó con el de neutral (ausencia de emoción). Se desarrolló un torso con brazos y manos para colocar el rostro robótico GolemX-1; una de las manos es móvil con 1 grado de libertad (1DOF). Además, se utilizó un sistema electromecánico como espacio de trabajo e interfaz de colaboración humano-robot(Figura 6.1).

La figura 6.2 muestra las áreas y elementos en la ambientación experimental; GolemX-1 se encuentra frente al sujeto en el lado opuesto de una



Figura 6.1: Prototipo facial GolemX-1 con torso, mano móvil e interfaz de colaboración

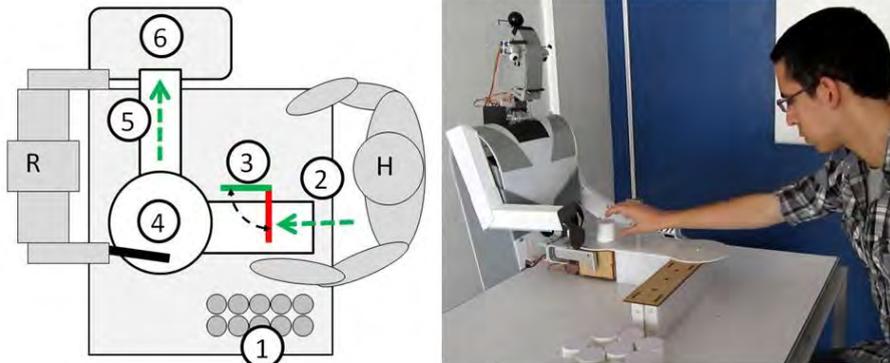


Figura 6.2: Espacio experimental de trabajo. (Izquierda) Sistema interactivo y espacio de trabajo colaborativo humano y robot. Diagrama: (1) Zona de objetos, (2) Acceso humano, (3) Barrera de acceso (rojo/cerrado, verde/abierto), (4) Área de lanzamiento y mano robótica, (5) Rampa, (6) Contenedor. (Derecha) Escenario y área de trabajo.

mesa. La mano derecha del robot está colocada sobre el área de lanzamiento, esperando a que la persona coloque el objeto para empujarlo hacia el contenedor. Se pidió al sujeto colocar un objeto a la vez. En el espacio de trabajo hay una barrera que regula la colocación de los objetos; el usuario solo podía colocar el objeto cuando la barrera estaba abierta y en caso contrario no podía intervenir en la acción.

El lanzamiento de los objetos o la interrupción de la tarea se realizó cuando la barrera se encontraba cerrada; el usuario se concretaba a observar la acción. El procedimiento experimental se realizó de la siguiente manera:

- a Se dan instrucciones escritas al usuario. El robot inicia con una cara

neutral.

- b El investigador realiza una prueba cómo muestra y coloca un objeto frente al robot; una vez colocado la barrera se cierra, el robot dirige la mirada al objeto y lo empuja hacia el contenedor. En el caso de la sesión de control, el robot mantiene el rostro neutral, de lo contrario, GolemX-1 activa el de felicidad.
- c El robot reinicia su posición con el rostro neutral en espera del siguiente objeto; la barrera se abre y GolemX-1 voltea el rostro para dirigir la mirada hacia la posición del usuario. El sujeto toma su lugar en la interacción y la actividad se reinicia.
- d El sujeto coloca el primer objeto frente al robot. En el primer intento el robot empuja exitosamente el objeto hacia el contenedor.
- e El humano coloca el siguiente objeto y se repite exitosamente el lanzamiento.
- f Las acciones del robot son aleatorias a partir del tercer lanzamiento en adelante; una omisión o supuesta falla en el lanzamiento del objeto puede ocurrir en cualquier momento. Durante la falla, el robot activa la expresión facial de acuerdo a la sesión (negativos o neutral). En el fallo, el robot continúa observando al objeto sin empujarlo hacia el contenedor. Se hace una pausa con 9 segundos de duración.
- g Después de la acción fallida, la barrera se abre por 12 segundos; el robot continúa observando al objeto con la cara neutral (durante

cualquiera de las tres sesiones).

- h Cuando los 12 segundos transcurren, la barrera se cierra y la acción fallida se repite. En el análisis, se prestó atención en la primera reacción o intervención humana que puede ocurrir durante la primera o segunda pausa. La falla se repite en dos ocasiones con el fin de estimular la acción humana.
- i Después de las dos fallas el robot empuja el objeto hacia el contenedor para continuar con la tarea. La falla no se repite.
- j La tarea se detiene cuando los diez objetos se encuentran en el contenedor.

Se asignaron diferentes tipos de retroalimentación emocional en tres sesiones distintas:

- a Sesión 1. El robot activa el rostro neutral durante toda la actividad (tarea de control).
- b Sesión 2. Se activa la expresión de felicidad durante los lanzamientos exitosos y de tristeza durante el fallo.
- c Sesión 3. Se activa la felicidad durante los lanzamientos exitosos y el de enojo durante el fallo.

Para el estudio se utilizaron dos cámaras de vídeo. El tiempo de ejecución en las tareas del robot fue asignado empíricamente durante sesiones previas con colegas. Los experimentos principales se llevaron a cabo con 15

sujetos por sesión; los cuales eran estudiantes de diseño industrial que no habían visto al robot con anterioridad. El promedio de edad de los sujetos fue de 22 años. Las acciones detalladas del experimento se especifican en el cuadro 6.1.

Cuadro 6.1: Descripción de la tarea.

Tarea	Acción humana	Acción del robot. Modo inexpressivo	Acción del robot. Modo expresivo
Inicio	Barrera abierta, el humano coloca el objeto frente al robot	Rostro neutral durante toda la tarea. El robot observa al objeto	Rostro neutral. El robot observa al objeto
Acción exitosa	Barrera cerrada el humano espera	El robot lanza el objeto y regresa la mano a la posición inicial	El robot activa rostro de felicidad, lanza el objeto y regresa la mano a la posición inicial
Finaliza acción exitosa	Barrera abierta, el humano puede poner el siguiente objeto	El robot voltea el rostro hacia el humano y espera el siguiente objeto	El robot activa rostro neutral, voltea el rostro hacia el humano y espera el siguiente objeto
Acción fallida	Barrera cerrada el humano espera	El robot omite el movimiento de la mano. Pausa de 9 segundos	El robot activa el rostro negativo, omite el movimiento de la mano. Pausa de 9 segundos
Finaliza acción fallida	Barrera abierta por 12 segundos. <b>Se espera la reacción del sujeto</b>	El robot observa al objeto	El robot activa el rostro neutral y observa al objeto

## 6.2. Resultados

Los resultados se obtuvieron después de revisar las vídeo grabaciones; se cronometró el tiempo de reacción de los usuarios y se identificaron cinco diferentes comportamientos en el usuario. Las acciones llevadas a cabo por los sujetos durante el supuesto fallo en el sistema se presentan a continuación:

1. El sujeto cambió la posición del objeto sobre el área de lanzamiento.
2. El sujeto cambió el objeto sobre el área de lanzamiento por otro de la zona de objetos.
3. El sujeto lanzó el objeto hacia el contenedor de la misma forma que el robot.
4. El sujeto colocó un objeto extra en el área de lanzamiento.
5. El sujeto esperó (inactividad).

Las tres primeras acciones tuvieron un efecto positivo; el humano ayuda al flujo de la tarea y realiza la actividad sin omitir las instrucciones. Sin embargo, las dos últimas intervenciones tienen un efecto negativo debido a que el sujeto no colaboró directamente al progreso de la tarea, se generaron sesiones más largas y no hubo continuidad con las reglas.

El cuadro 6.2 muestra el tipo de comportamientos y sus frecuencias. Los resultados muestran que las expresiones negativas tienen un efecto positivo en la tarea.

Cuadro 6.2: Tipo de comportamientos en la tarea y retroalimentación

Acción-intervención	Neutral	Tristeza	Enojo
Cambio en la posición del objeto	7	14	14
Cambio de objeto	1	1	0
Lanzamiento de objeto	0	0	1
Poner más objetos	4	0	0
Esperar	3	0	0

Las frecuencias de los efectos positivos y negativos durante las sesiones se presentan en el cuadro 6.3. La expresión de tristeza tiene una diferencia estadística cuando se compara con el de neutral ( $p \leq 0,05^1$ ). El efecto de enojo tiene una diferencia estadística similar cuando se compara con el de neutral; tristeza y enojo no son diferentes estadísticamente ( $p \geq 0,05^2$ ).

Los promedios en los tiempos de reacción registrados (intervención) fueron de: 21 segundos para neutral, 12 para tristeza y 7 a favor de enojo. Hay una evidencia empírica en la cual el enojo del robot incita al usuario a reaccionar más rápido. Estadísticamente no fue posible afirmar alguna diferencia en los tres casos ( $p \geq 0,05$ ).

<sup>1</sup>Análisis completo en <https://goo.gl/pGE16K>.

<sup>2</sup>La significación estadística fue calculada por el método de la prueba de "T" pareada (Pair t-test) para un experimento con potencia de 0,9 y muestra de 15 por caso.

Cuadro 6.3: Efecto positivo y negativo durante las sesiones

Efecto en la tarea	Expresión		
	Neutral	Tristeza	Enojo
<b>Efecto Positivo</b>	8	15	15
<b>Efecto negativo</b>	7	0	0

### 6.3. Discusión

Los rostros negativos de tristeza y enojo obtuvieron puntuaciones altas en las pruebas de legibilidad por encima de los positivos; en particular el de enojo obtuvo el resultado más alto. Se analizó la influencia de las expresiones robóticas negativas en comparación con el neutral (ausencia de emoción); básicamente, se estudió el efecto de la emoción negativa como retroalimentación durante una tarea colaborativa humano-robot. Durante el experimento los negativos fueron activados (uno por sesión) cuando el robot es incapaz de arrojar objetos a un contenedor; la expresión emocional sirvió para mostrar una señal no verbal que advertía de una situación adversa al progreso de la tarea. En esencia, la falla repentina representa la causa de la activación del gesto negativo que el sujeto observa y reacciona en consecuencia.

Los resultados señalaron que las emociones negativas tienen un efecto regulador en la interacción; los gestos orientan a los usuarios para completar una tarea, seguir las instrucciones puntualmente y mantener adecuadamente el flujo de la tarea. Hipotéticamente, tal efecto se favorece por la

relevancia de las expresiones negativas en la atención humana y la función de estos para comunicar situaciones de advertencia (enojo) e incapacidad (tristeza).

La ausencia de expresiones durante la actividad (neutral) demostró ser una retroalimentación incierta en relación al estado de la tarea; algunos sujetos no cooperan con el avance y reaccionan con menor precisión al omitir las instrucciones dadas. Considerando que tristeza y enojo resultaron estadísticamente similares, es posible afirmar que el efecto positivo es independiente de la excitación alta en enojo y baja en la tristeza según los modelos dimensionales de emoción (Modelo de Activación positiva - Activación negativa [53, 54] y Modelo Circumplejo de Afecto [5, 50–52]). Por ejemplo, cuando hay un incremento en el nivel de excitación y hay un cambio de tristeza a enojo; ambas expresiones difieren en la condición que las detona y su función dentro de la interacción social [2]. Además, su contraparte positiva dentro de los modelos dimensionales de emoción tiene un efecto similar respecto a su nivel de excitación; la valencia positiva para la emoción de tristeza es alegría y para el de enojo es el estado relajado o contento (buen humor) que no son consideradas como emociones universales.

La versión contraria de las emociones puede ser útil para enfatizar una emoción dentro del contexto; una expresión negativa puede ser mejor identificada si se activa durante la presencia de su contraparte positiva, en tal caso, el efecto de la expresión facial sobre el observador es un tema a investigar.

Este experimento se reportó en un artículo y ponencia denominada “The positive effect of negative feedback in HRI using a facial expression robot” presentado en el Taller Internacional de Robótica Cultural (International Workshop in Cultural Robotics) dentro del Simposio Internacional “The 24th International Symposium on Robot and Human Interactive Communication” llevado a cabo en Kobe, Japón durante 2015. Para mayor detalle referirse a [49]

## Capítulo 7

# La expresión robótica de enojo en interacción humano robot (HRI)

La expresión facial de enojo de GolemX-1 fue la expresión facial más legible durante las pruebas de identificación, tiene influencia positiva cuando se utiliza como retroalimentación durante un fallo además de afectar la precisión y reacción del usuario ya que reduce el tiempo de intervención durante la actividad. La expresión de enojo fue susceptible a un tercer análisis en donde se evaluó el efecto del rostro sobre la atención humana. La literatura que estudia el efecto de enojo en el entorno social, destaca su utilidad para hacer entendibles circunstancias poco claras durante la interacción sin ser comunicadas verbalmente [177].

## 7.1. Antecedentes

Las expresiones faciales ayudan a comunicar el estado emocional de los individuos; se espera que su efecto sea consistente en quienes las observan. Un problema durante la interacción social es la incertidumbre que genera el desconocimiento del estado mental de las personas y por ende su proceder [178]. El observador interpreta el estado mental de otros pues es imposible percibirlo directamente [179]. La activación de expresiones faciales tiene una función social predeterminada a partir de ciertas condiciones que la detonan y funciones en el entorno [2,9,180]. Las expresiones emocionales pueden ser una ventana al estado mental de los individuos en relación al entorno de interacción; este efecto ha sido utilizado en agentes artificiales con el fin de mostrar un estado emocional y persuadir al usuario [169].

Las caras negativas son mejor reconocidas debido a su utilidad para informar acerca de una potencial amenaza a miembros de un grupo social [25]. El enojo es una emoción asociada al odio, error, agresión y contrariedad [24]. La hipótesis que aborda su estudio está relacionada con el decremento del auto-control emocional e incremento en la irritabilidad, agresión e insatisfacción [23]. La función biológica de la expresión evolucionó y se adaptó al entorno social; puede utilizarse para obtener beneficios a favor de quien la expresa dentro de entornos poco claros y desordenados [32].

El enojo puede ser útil para dirigir la interacción entre individuos y puede activarse cuando las metas u objetivos no se cumplen; es señal de que algo no está bien y tiene que ser modificado [2]. El enojo activa un proceso

de análisis y diagnóstico en el observador [177]. Con el fin de estudiar tal efecto durante la interacción humano robot se implementaron expresiones negativas en rostros artificiales [30, 57].

La teoría de la recalibración del enojo (The Recalibrational Theory of Anger) asume que el sistema regulador que dirige la emoción se convirtió en un medio para negociar y resolver conflictos de interés además de recalibrar la situación para el bienestar del sujeto que la expresa [32, 33, 181]. En general, la cara robótica de enojo es relativamente reconocida con una buena precisión [14]; esto a servido para tomar ventaja en tareas de robots de servicio e interactivos (Robot Minerva y robot iCat [30, 57]).

A partir de los efectos documentados de enojo y al efecto positivo dentro de una actividad colaborativa de acuerdo al experimento anterior, se analizó su influencia durante una tarea. Con el fin de generar mayor atención hacia la expresión facial fue necesario considerar el contexto de activación. Básicamente se analizó el efecto de la cara de enojo para inducir la atención humana durante una actividad poco legible.

## **7.2. El efecto del enojo sobre la atención humana**

El experimento para evaluar la atención humana fue diseñado con algunos elementos complementarios para la presentación de expresiones emocionales; en particular el uso de expresiones transicionales o dinámicas. Comúnmente, el cambio entre expresiones faciales se presenta a través de transiciones de un estado emocional a otro con diferente intensidad y va-

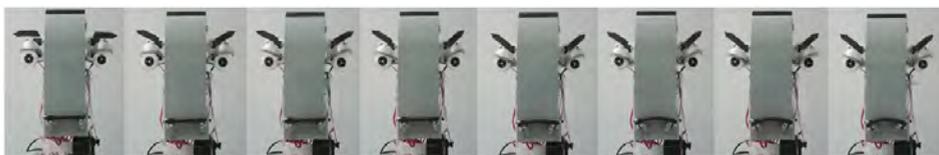


Figura 7.1: Conjunto transitivo negativo. Neutral a enojo con excitación alta (Derecho a izquierdo, neutral a intensidad alta).

lencia. Una expresión transicional difiere de una estática en que es una secuencia de expresiones con intensidad incremental [107,108]. Las figuras 7.1 y 7.2 muestran ejemplos de expresiones dinámicas con GolemX-1 (neutral a enojo y de neutral a felicidad); con el fin de alcanzar una intensidad alta en la emoción de felicidad, el rostro fue modificado en la boca y los ojos del robot para expresar una sonrisa de tipo Duchenne [182]. Este rostro es una variante de la expresión positiva pero catalogada como auténtica, genuina y confiable [183] la cual se activa de forma espontánea ante un estímulo inesperado; a diferencia de una expresión de felicidad común, el cuál puede ser expresado conscientemente como cortesía, la sonrisa Duchenne es producto de una reacción súbita positiva. La meta de la transición fue simular una excitación alta de acuerdo a una situación en la cual se muestra mayor felicidad o enojo.

La presencia súbita de una expresión facial simple (estática) puede ocurrir durante el contexto de los dinámicos; una expresión positiva transicional puede contextualizar la presencia de un estático negativo a partir del cambio en la valencia y la interrupción en la secuencia de transición. Con el fin de estudiar la efectividad de enojo para inducir la atención humana,

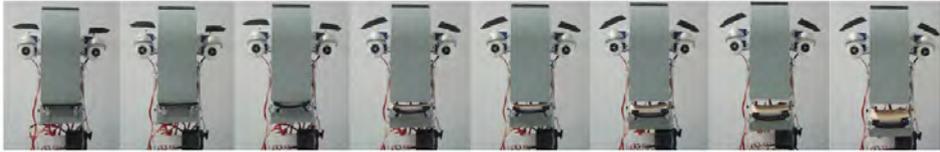


Figura 7.2: Conjunto transitivo positivo. Neutral a feliz con excitación alta (Derecho a izquierdo, neutral a intensidad alta).

las emociones de sorpresa y felicidad (positivos) se presentaron en conjunto durante una actividad colaborativa humano y robot. La figura 7.3 muestra tres situaciones de la activación súbita de un gesto estático durante una expresión transicional. En el diagrama, la situación **A** muestra la ejecución de enojo durante la transición de neutral a felicidad. La situación **B**, la expresión de felicidad durante la transición de neutral a enojo y la situación **C**, uno de sorpresa durante la transición de neutral a enojo.

El escenario experimental diseñado presenta una variación en los rostros de transición y estáticos, estos pueden ocurrir en secuencia, uno seguido del otro, aunque imposible de predecir. La expresión de la emoción estática ocurre de forma arbitraria dentro de la interacción y contexto de la expresión transitiva. Esta situación fue evaluada dentro del escenario del capítulo anterior aumentando el tipo y la cantidad de objetos; la figura 7.4 muestra el espacio de experimentación. El humano y el robot cooperan para colocar objetos cilíndricos dentro de un contenedor; en esta ocasión se incluyen 4 columnas con 4 objetos de color azul, amarillo, verde y blanco (los objetos en cada columna tienen el mismo color). Las columnas fueron colocadas aleatoriamente en cada sesión. El sujeto elige cualquier objeto

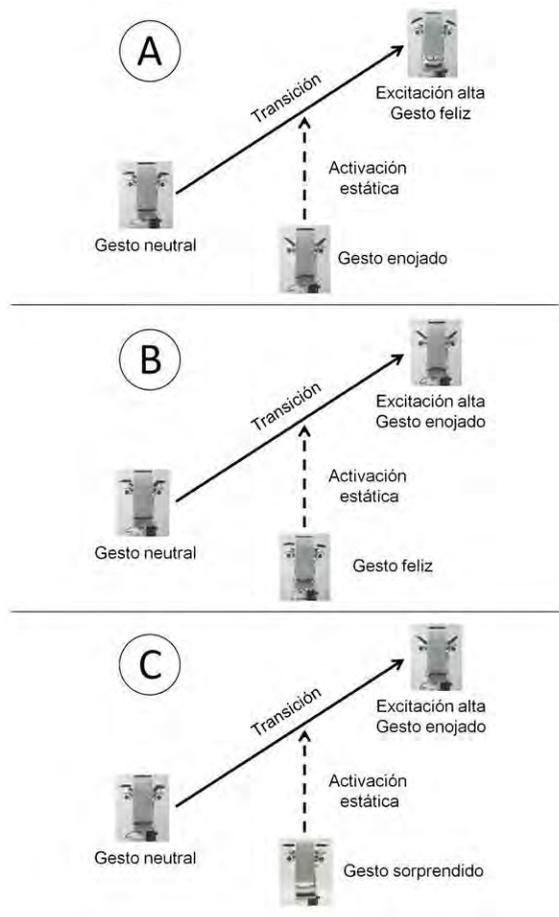


Figura 7.3: Contexto de expresiones faciales transitivas y la activación súbita de una expresión estática. De Izquierda a derecha, Situación **A**: Estático negativo y Transitivo positivo. Situación **B**: Estático positivo (feliz) y Transitivo negativo. Situación **C**: Estático positivo (Sorpresa) y Transitivo negativo.

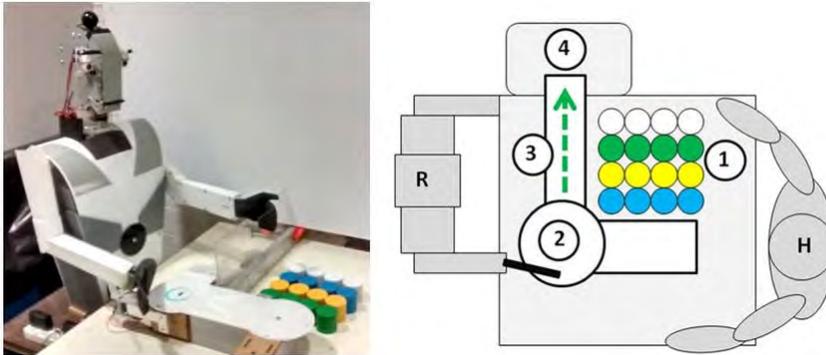


Figura 7.4: Escenario de experimentación. Sistema de interacción y diagrama de trabajo colaborativo humano y robot: (1) Zona de objetos, (2) Área de lanzamiento y mano robótica, (3) Rampa, (4) Contenedor.

para colocarlo en el área de lanzamiento y el robot lo empuja hacia al contenedor. Al final de la sesión, se aplicó un cuestionario para evaluar el desempeño del robot, la legibilidad de la expresión y la causa en la activación.

Golem X-1 activa un rostro en transición cuando los objetos amarillos, verdes y blancos se encuentran en el área de lanzamiento. Con el fin de establecer una situación confusa o no entendible, el robot activa uno estático cuando se presenta el color azul. El sujeto no está informado acerca de la activación de rostros. En caso de que el humano no coloque el objeto, el robot activa la expresión facial estática. En todos los casos el robot realiza el movimiento de la mano y lanzamiento del objeto. Para la activación del estático, se eligió el objeto del color azul ya que no está relacionado con algunas de las emociones utilizadas; el color azul es interpretado como tris-

teza [184, 185] (por ejemplo, la metáfora de “Me siento azul“ o “i feel blue methaphor“) y es utilizado en caras robóticas [186].

El experimento se realizó de la siguiente manera:

- a) Se dan instrucciones escritas al usuario y el investigador realiza una sesión de prueba.
- b) El investigador reinicia el sistema.
- c) El sujeto coloca el objeto.
- d) El robot mira al objeto, lo empuja hacia el contenedor y el rostro se activa (el ciclo completo dura aproximadamente 3 segundos).
- e) El robot espera al siguiente objeto (durante aproximadamente 4 segundos).
- f) La acción se repite hasta que todos los objetos se colocan dentro del contenedor.
- g) La tarea termina y el sujeto responde el cuestionario.

### **7.3. Cuestionario**

El cuestionario se presenta en el cuadro 7.1. La primera pregunta cuestionó acerca de la facilidad de la tarea; si fue percibida como muy difícil a muy fácil. La siguiente, acerca de la velocidad del robot para realizar la tarea de muy lento a muy rápido. En las preguntas 1 y 2 lo sujetos eligieron

Cuadro 7.1: Cuestionario del experimento colaborativo humano-robot

Preguntas	Opciones / Respuestas
1.¿Fue fácil la tarea?	Muy difícil / Poco difícil /Normal/ Poco fácil/ Muy fácil
2.¿Qué tan rápido trabajó el robot?	Muy lento/ Poco lento / Normal/ Poco rápido/ Muy rápido
3.¿Lograste identificar las expresiones en el robot?	Si / No
4.¿Qué expresiones pudiste identificar?	Feliz/ Enojo / Sorpresa/ Miedo/Tristeza
5.¿Cuáles crees que fueron las causas de la activación del gesto?	Pregunta abierta

1 de 5 opciones de una escala de Likert con valores de 0 a 100 puntos (0, 25, 50, 75 y 100).

Las preguntas 3 y 4 se implementaron para medir la legibilidad de las expresiones faciales estáticas y transicionales; la pregunta 3 evaluaba si los sujetos identificaron las emociones durante la interacción. Se utilizó una escala de 0 (no) o 100 (si) puntos. La pregunta 4 fue abierta y cuestionaba por los rostros que el sujeto pudo identificar en la cara robótica; se dieron cinco opciones (Felicidad, enojo, sorpresa, miedo y tristeza).

La pregunta 5 fue abierta e indagó acerca de las posibles causas que activaron las expresiones faciales durante la tarea; se puso atención sobre la descripción exacta de la causa la cuál podía ser válida sin importar que el

sujeto utilizara una descripción diferente, por ejemplo: “El rostro se activó cuando el objeto azul está en el area”, “El objeto azul”, “La emoción se activó porque al robot le gusta / disgusta el objeto azul”, “ La expresión se activó cuando los objetos amarillos, verdes y blancos estaban en el área de lanzamiento”, “Todos los objetos excepto el azul”, “ El gesto se activó durante la tarea”, etc.

## 7.4. Resultados

El experimento se realizó con 69 participantes (23 sujetos no expertos por escenario). Todos los sujetos eran colegas y estudiantes de licenciatura reclutados por invitación personal. El cuadro 7.2 muestra los resultados de las primeras tres preguntas. La primera columna muestra que la mayoría de los usuarios consideraron a la tarea cómo fácil; lo cual corresponde al diseño de una tarea no compleja y repetitiva. La segunda muestra que los sujetos afirmaron que la velocidad de la tarea fue normal y la tercera que todos los sujetos identificaron expresiones durante el experimento.

Los resultados de la identificación de emociones se presentan en el cuadro 7.3. En las situaciones **A**, **B** y **C** la expresión transicional fue reconocida con una precisión alta (91,3 %, 100 % y 100 % respectivamente). Los estáticos también fueron bien reconocidos en las tres situaciones (100 %, 95,7 % 95,7 % respectivamente). Sin embargo, hubo una confusión entre sorpresa y felicidad en las situaciones **A** y **C**. En la primera, sorpresa fue reconocido, aunque no se activó durante la sesión (78,3 %). Hipotéticamente, la activa-

Cuadro 7.2: Cuantificación de las respuestas 1 a la 3

<b>Preguntas</b>	<b>Sesión A</b>	<b>Sesión B</b>	<b>Sesión C</b>
<b>Pregunta 1.</b>	94,26 %	89,13 %	95,65 %
<b>Pregunta 2.</b>	58,69 %	57,60 %	56,52 %
<b>Pregunta 3.</b>	100,0 %	100,0 %	100,0 %

ción de la boca y ojos durante la intensidad alta de la expresión facial de felicidad (“Duchenne smile”) fue confusa para el observador. En la última situación, felicidad fue identificada en algún momento de la tarea (56,5 %). Esta confusión no fue consistente en todas las situaciones; en **B** felicidad fue mejor identificado que sorpresa (39,1 %) y en **C** sorpresa obtuvo una puntuación más alta que felicidad. Aunque las expresiones de miedo y tristeza no se activaron, algunos sujetos las registraron, pero con un resultado muy bajo.

El registro de las causas de la activación de emociones (con código de identificación) se muestra en el cuadro 7.4. Todas las causas se esperaban como resultado válido durante el diseño del experimento excepto por **DP** (Retraso en colocar el objeto) el cual también fue considerado como válido. Causas diferentes o imprecisas no fueron consideradas. Los participantes pudieron identificar una o dos causas, pero en algunos casos no lograron identificar alguna.

En la figura 7.5 se compara el número de causas identificadas por sesión (0 a 2 causas) y número de sujetos durante expresiones estáticas. En

Cuadro 7.3: Resultados de la pregunta 4. Registros de las expresiones faciales.

	<b>Situación A</b>	<b>Situación B</b>	<b>Situación C</b>
Enojo	<b>Estático</b> 100 %	<b>Transicional</b> 100 %	<b>Transicional</b> 100 %
Felicidad	<b>Transicional</b> 91.3 %	<b>Estático</b> 95.7 %	Inactivo 56.5 %
Sorpresa	Inactivo 78.3 %	Inactivo 39.1 %	<b>Estático</b> 95.7 %
Miedo	Inactivo 8.7 %	Inactivo 8.7 %	Inactivo 4.3 %
Tristeza	Inactivo 17.4 %	Inactivo 30.4 %	Inactivo 17.4 %

Cuadro 7.4: Causas de la activación del gesto y su codificación

<b>Causas</b>	<b>ID/Código</b>
<b>Activación de gestos transicionales</b>	
1. Progreso de la tarea	PT
2. Todos los objetos excepto el azul	AOxB
<b>Activación de gestos estáticos</b>	
1. El objeto azul en el área de lanzamiento	BO
2. No hay objetos en el área de lanzamientos	NO
3. Retraso en colocar los objetos	DP

la figura 7.6 se muestran los resultados incluyendo las sesiones completas (estáticos y transicionales). Los resultados se corroboraron con análisis estadístico; la situación **A** es significativamente diferente de **B** y **C** ( $p < 0,05$ ), mientras que las últimas dos no son diferentes entre si ( $p \not< 0,05$ <sup>1</sup>). El enojo induce a la atención humana mejor que felicidad y sorpresa, especialmente cuando es activado súbitamente entre los transicionales.

Los sujetos identificaron con precisión alta las causas que activan los estáticos de enojo; aunque durante su activación transicional no fueron fáciles de identificar. Hipotéticamente, la presencia prolongada de enojo confunde al observador y hace menos legible la activación de los gestos estáticos positivos.

<sup>1</sup>El análisis completo se encuentra en <https://goo.gl/pGE16K>

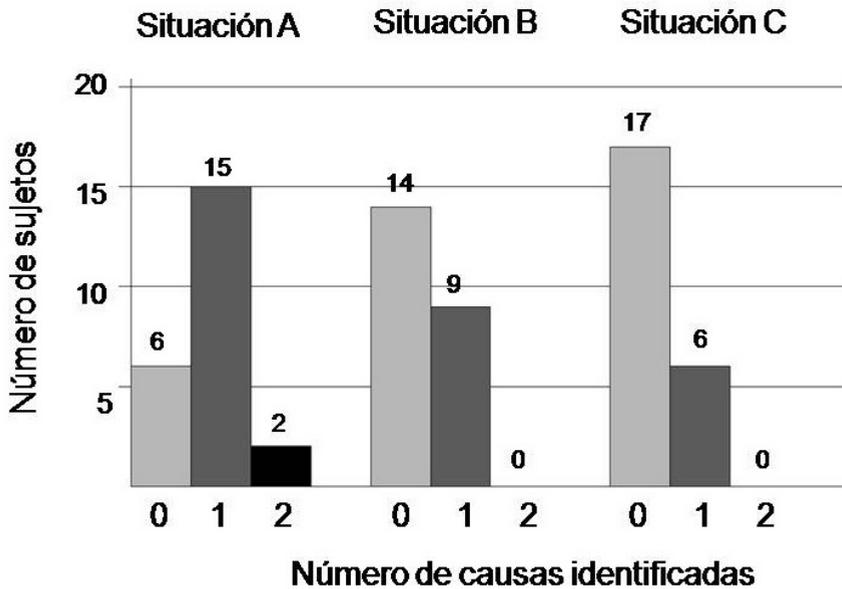


Figura 7.5: Sesiones y número de causas identificadas durante la activación de la expresión estática.

En el cuadro 7.5 se presentan el número de causas identificadas durante las sesiones. El cuadro 7.6 muestra el número de sujetos y causas por cada situación durante la activación de estáticos; los resultados para la activación de expresiones estáticas y transicionales se presentan en el cuadro 7.7.

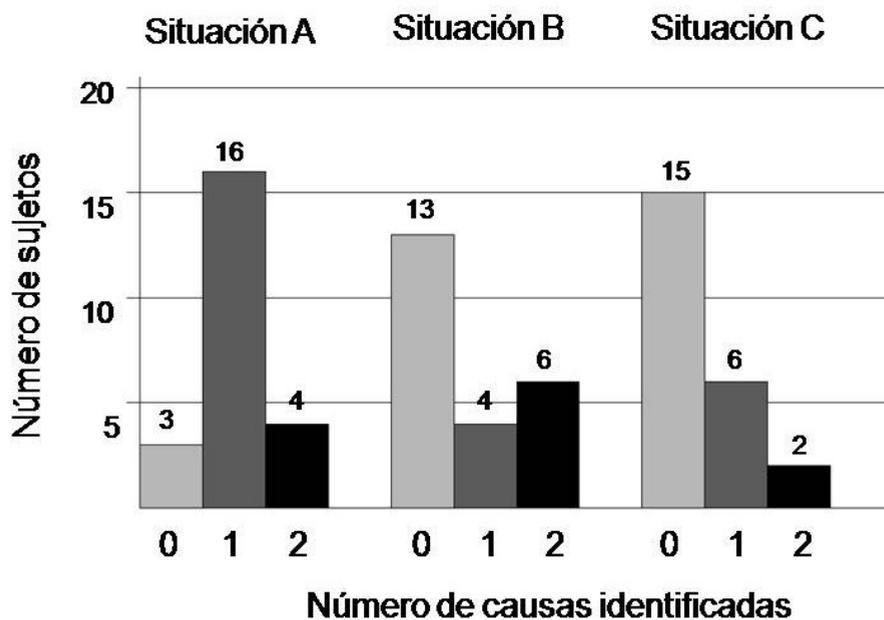


Figura 7.6: Sesiones y número de causas identificadas durante la activación de expresiones estáticas y transicionales.

Cuadro 7.5: Causas registradas de la activación de expresiones faciales

<b>Emoción</b>	<b>Situación A</b>	<b>Situación B</b>	<b>Situación C</b>
<b>Enojo</b>	<b>Estático</b>	<b>Transicional</b>	<b>Transicional</b>
	Causas Registro	Causas Registro	Causas Registro
	BO 11	AOxB 7	PT 1
	NO 3		AOxB 3
	DP 5		
<b>Feliz</b>	<b>Transicional</b>	<b>Estático</b>	<b>Inactivo</b>
	PT 5	BO 9	
<b>Sorpresa</b>	<b>Inactivo</b>	<b>Inactivo</b>	<b>Estático</b>
			BO 5
			NO 1

Cuadro 7.6: Numero de causas por situación durante las expresiones estáticas

Número de causas identificadas	Número de sujetos y causas		
	A	B	C
0	6	14	17
1	9 (BO) 5 (DP) 1 (NO)	9 (BO)	5 (BO) 1 (NO)
2	2 (BO+NO)	0	0

Cuadro 7.7: Numero de causas por situación durante las expresiones estáticas y transicionales

Número de causas identificadas	Número de sujetos y causas		
	A	B	C
0	3	13	15
1	3 (PT) 9 (BO) 3 (DP) 1 (NO)	1 (AOxB) 3 (BO)	3 (BO) 1 (NO) 2 (AOxB)
2	2 (BO+NO) 2 (PT+DP)	6 (BO+AOxB)	1 (BO+PT) 1 (BO+AOxB)

## 7.5. Discusión

La cabeza robótica Golem X-1 es capaz de comunicar cinco de seis expresiones universales. Después de los análisis de legibilidad en la expresión facial, se encontró que enojo es la mejor identificada. El resultado dio la pauta para analizar la influencia del gesto en los humanos dentro de una actividad colaborativa (humano y robot); el rostro fue comparado con el de neutral y tristeza. La expresión emocional fue utilizada como retroalimentación en una tarea simple y repetitiva. Durante la tarea, el rostro de enojo fue activado durante un fallo intencionado y programado en el sistema lo cual tuvo un efecto positivo para regular las acciones humanas a favor de la secuencia y precisión [49]; este hallazgo fue consistente con la expresión facial de tristeza.

En el último experimento, una secuencia de expresiones en el robot se asignó al progreso de una tarea. Durante la secuencia se mostró una transición de rostros que representa el incremento en la intensidad de una emoción que iba de neutral a la máxima intensidad emocional que pudo simularse en la cara robótica. Los transicionales fueron usados para crear un contexto de interpretación y activación de los estáticos.

La transición diseñada para el contexto de los estáticos está compuesta por un rostro neutral a felicidad con intensidad alta y neutral a enojo. La presencia súbita de la expresión estática destaca un cambio en el contexto; la valencia en las emociones es la base de tal cambio. Es posible realizar un contraste en las transiciones a partir del cambio de emociones positivas a negativas y vice versa.

El estudio muestra que la presencia súbita del enojo en el contexto de una expresión transicional positiva resulta efectiva para inducir la atención en sujetos no expertos; un considerable número de participantes registró al menos una causa relacionada con la activación de la expresión estática bajo condiciones no explícitas y confusas durante la tarea. El enojo promueve el análisis y la evaluación de una situación durante condiciones poco claras; el sujeto que observa la emoción indica con mayor efectividad la causa de su activación. Sin embargo, la presencia continua del rostro negativo es adverso a su legibilidad y su significado dentro de la interacción.

Finalmente, varios sujetos infirieron una preferencia del robot hacia los objetos cuando argumentaron un gusto o disgusto producto de las emociones de felicidad y enojo. En este caso el sujeto dedujo un supuesto estado mental en el robot a partir de la emoción, la situación y el entorno. La activación de expresiones faciales puede estar relacionada a una intención demostrada a través del comportamiento y postura, específicamente, al mostrar una preferencia o rechazo, interés o desinterés, acuerdo o desacuerdo, etcétera.

## Capítulo 8

# La activación de expresiones emocionales en Golem-III

Los robots de servicio se utilizan para asistir a las personas dentro de espacios sociales tales como el hogar, oficina, tiendas, hospitales, museos, etcétera. En el futuro, tales equipos necesitarán de sistemas flexibles para operar de forma eficiente en diversos escenarios; un robot con estas características deberá mostrar un comportamiento perceptivo capaz de inferir ante un estímulo y acción en particular. El desarrollo de sistemas autónomos toma en cuenta el contexto de uso para desplegar funciones predefinidas o bien comportamientos reactivos que regulen las acciones y adaptarse a una situación emergente.

La interacción entre personas y robots involucra fenómenos que afectan de diversas maneras la operación de los sistemas. Es necesaria una base de

datos capaz de responder a las acciones de otros agentes o eventos externos impredecibles que ocurren durante la interacción; todo esto, sin alterar el cumplimiento de las tareas encomendadas.

En el caso de las personas, el comportamiento que regula la respuesta ante fenómenos emergentes es muy complejo; para emularlo, se ha simplificado a través de protocolos flexibles ejecutables durante las actividades del robot tales como tareas prácticas [20]. Así, el sistema se encuentra situado en el contexto, es capaz de ejecutar y concluir una tarea práctica cumpliendo con actividades parciales (situaciones) para completar otras de mayor complejidad. Esta interacción se modela en el robot Golem-III a través de la Arquitectura Cognitiva Orientada a la Interacción [21, 22] descrita en el capítulo 4

## 8.1. Inferencia oportunista

En general, los robots de servicio utilizan arquitecturas cognitivas para desempeñarse dentro de ambientes reales (complejos) por lo que su desarrollo deberá exhibir un comportamiento adaptable. Los esquemas de funcionamiento procedimental en robots pueden tener cierta estabilidad para ejecutar sus acciones, sin embargo, la omisión de eventos esperados por el sistema puede obstaculizar el logro de objetivos e incluso provocar que la tarea no se cumpla.

Con el fin de crear ciclos de interacción consistentes se ha integrado en la operatividad del robot Golem-III, representaciones del estado externo a

través de expectativas acerca de las intenciones de los interlocutores o eventos naturales relevantes que puedan ocurrir en el entorno de tal forma que se emule un comportamiento similar al pensar y actuar. Las expectativas en el robot son producto del procesamiento de la información detectada en el mundo sujeta a la comparación con la representación interna del sistema, lo cual produce un comportamiento intencional para completar o alterar la situación de acuerdo al estado óptimo de la misma. Las acciones externas del robot se muestran a través de conductas motoras y de lenguaje, en el caso de las internas, se realizan a través de inferencias deductivas [187]. El robot estará situado si su representación del exterior es coherente con el estado del mundo.

El comportamiento del robot puede estar modelado a partir de la representación explícita de la tarea la cual se complementa con la del contexto de ejecución; el robot alcanzará progresivamente las metas conforme al cumplimiento de los escenarios específicos en la estructura de la tarea. La inferencia puede utilizarse a petición en el caso de que las expectativas en el sistema no se cumplan; el agente invoca los patrones de análisis y evaluación hasta que determina las acciones a seguir para conseguir los objetivos. En este caso, una vez que el robot reconfigura las acciones para completarlas es capaz de continuar la tarea utilizando los patrones de comportamiento básicos (inferencia oportunista) [188].

La inferencia ha sido utilizada con diferentes formatos en robots de servicio, desde sistemas que utilizan planeadores que deciden las acciones del robot a partir de los comandos del usuario [189], sistemas que utili-

zan métodos similares dentro de arquitecturas cognitivas para la planeación de tareas probados en competencias internacionales (RoboCup, RoboCup@Home) [190] y que integran el razonamiento cualitativo durante la interacción humano-robot. También se han desarrollado varios sistemas que incorporan el procesamiento de la información, toma de decisiones y generación de planes para tareas complejas de manipulación [191]. Para la toma de decisiones en robots sociales se desarrolló una arquitectura deliberativa en la cual el sistema planifica las acciones durante tareas compartidas con humanos; el sistema toma en cuenta las intenciones, creencias y habilidades del colaborador humano [192]. Con la intención de generar conductas en robots que permiten aprender a tomar decisiones se desarrolló un sistema interactivo que planifica sus acciones a partir de instrucciones dadas por el operador [193].

El ciclo de inferencia desarrollado por el grupo Golem involucra tres procesos: a) Diagnóstico. El robot realiza una hipótesis con el fin de inducir las condiciones que difieren entre su concepto y el estado externo del mundo. b) Toma de decisiones. El sistema realiza una inferencia cuya finalidad es decidir qué hacer. c) Planificación. El robot infiere y organiza la ejecución de acciones para completar la tarea. Los procesos se han investigado de forma aislada en estudios de inteligencia artificial, sin embargo, el ciclo de inferencia en Golem-III se realiza de forma conjunta y sistemática. Este enfoque es útil para generar un comportamiento robusto en robots de servicio con el fin de solucionar problemas complejos y emergentes.

El concepto de situaciones permite representar la estructura de la tarea

de forma similar a un gráfico con las actividades que el robot necesita desplegar para alcanzar su objetivo. Esta estructura puede considerarse como un protocolo que configura el comportamiento de un agente artificial capaz de realizar tareas de servicio en contextos sociales. Golem-III utiliza el módulo de inferencia para realizar este tipo de tareas de forma flexible y confiable [188]. El módulo es parte de la arquitectura cognitiva del robot y coordina la percepción, la inferencia y la acción dentro del ciclo de comunicación e interacción. En [188] se muestra un caso de estudio que incorpora esta funcionalidad en Golem-III durante una tarea; el robot asume el papel de asistente en un supermercado.

La estructura de la tarea permite una representación emocional simple a través de los gestos emocionales en Golem-III; es posible asignar una etiqueta emocional apropiada cuando se cumple alguna expectativa durante la tarea. Así, la inferencia en el sistema facilita la implementación simple de un modelo del comportamiento emocional sujeto a las expectativas, logros y situaciones inesperadas durante la interacción. El robot Golem-III es capaz de mostrar cinco de seis expresiones universales (excepto disgusto). El comportamiento emocional en el robot es regulado por la arquitectura cognitiva (IOCA) [21].

La expresión de emociones se ha utilizado en robots de servicio para mejorar la interacción con humanos (por ejemplo, el robot Lisa [88] y el robot Bender [105]). Varios robots expresivos utilizan emociones universales de acuerdo a los propuestos por Ekman (Felicidad, tristeza, sorpresa, enojo, miedo y disgusto) [37]. Se ha demostrado que los sistemas robóticos expre-

sivos favorecen la retroalimentación y enlace con los humanos [10, 18, 19].

Las expresiones faciales sirven para comunicar el estado mental de los humanos el cuál se espera que su efecto y expectativas sea consistente; pueden interpretarse como una ventana al estado mental de las personas y en el caso de robots mostrar de forma no verbal las condiciones internas del sistema. Las expresiones pueden ser útiles para hacer explícitas las situaciones poco claras que surgen durante la interacción.

El comportamiento emocional de un sistema autónomo interactivo dentro de un ciclo de inferencia reduce la incertidumbre respecto a su estado interno. Como se ha revisado hasta ahora, la expresión facial puede mostrar un comportamiento emocional inteligente y motivación individual; hace que la conducta sea comprensible y predecible. Además, regula, sincroniza, organiza y completa la interacción [13].

Las expresiones emocionales activadas de forma oportuna proporcionan información adicional a la situación a partir del estado interno del robot. Durante la interacción es posible que se presenten situaciones confusas o poco evidentes para el usuario, en este caso, el uso de gestos emocionales pueden orientarlo e incluso sugerir la intervención cooperativa para contribuir al alcance de objetivos. Como se ha analizado a lo largo de esta tesis, la valencia en las emociones informa de forma práctica el estado de una tarea, ya sea para determinar su éxito o fracaso e incluso persuadir al usuario a realizar un diagnóstico e inferir respecto a una situación sin el uso de señales sociales verbales. La activación de las expresiones faciales en Golem-III es el resultado del diagnóstico y una supuesta decisión en el

sistema, además, una emoción en la cara robótica puede informar cuando se inicia el ciclo de inferencia. Las expresiones faciales y las situaciones en las que se expresan hasta el momento, son las siguientes:

- a) Feliz: El objetivo se completó o las condiciones son consistentes con el estado hipotético del mundo.
- b) Sorpresa: El mundo no es el esperado en la situación (es decir, no se cumple ninguna expectativa en el mundo).
- c) Tristeza: Las expectativas para completar el objetivo actual son bajas y la inferencia de diagnóstico debe comenzar.
- d) Enojado: Una condición adversa todavía está presente en el ambiente y el ciclo de inferencias de la vida diaria debe continuarse. El gesto induce a la participación del usuario si las condiciones favorecen a su colaboración, análisis, diagnóstico e intervención.
- e) Miedo: El sistema no puede completar la tarea y se debe invocar un protocolo de recuperación de ayuda.

En la página: <https://www.youtube.com/watch?v=7kA2pir6vrA>, se muestra un vídeo de Golem-III realizando una tarea como asistente de supermercado. Durante la actividad se presentan algunas inconsistencias en el contexto de interacción las cuales el robot hace uso del ciclo de inferencia para cumplir exitosamente con los objetivos.

## 8.2. Sumario

La mayoría de los comportamientos emocionales en robots interactivos se realizan a través de arquitecturas basadas en sistemas cognitivos que evalúan los fenómenos externos y configuran un estado emocional como respuesta. Esta reacción muestra un supuesto estado mental y postura del robot. Es deseable que la representación emocional tenga sentido en el contexto de la actividad y sea significativa para el humano que interactúa, además de que influya positivamente en la interacción.

Muchos de los sistemas emocionales cognitivos asignan valores incrementales a situaciones predeterminadas en el entorno; en robots sociales, esta asignación puede realizarse a criterio del investigador de forma significativa y entendible para el usuario. Es deseable que el comportamiento emocional de los robots de servicio esté ligado a la tarea del sistema y a las situaciones en el entorno.

Las tareas de los robots de servicio determinan el comportamiento general del sistema. La integración lógica y estructurada de sus funciones son fundamentales para el logro de objetivos, sin embargo, durante situaciones del mundo real se presentan condiciones poco favorables al desempeño del robot; un sistema flexible capaz de reestructurar la secuencia de la tarea incrementa las posibilidades de concluir la actividad exitosamente.

La capacidad de un sistema para modificar sus acciones a partir de las condiciones no esperadas en el entorno representa un reto de la inteligencia artificial; estas condiciones resultan mucho más complejas si existe la intervención humana.

La arquitectura cognitiva del robot Golem-III se complementa con un ciclo de inferencia que involucra tres procesos: diagnóstico, decisión y planificación. El ciclo se realiza a partir de la suposición del robot acerca del estado del mundo; en caso de existir una diferencia entre la expectativa y la situación real, se inicia el ciclo de inferencia de para lograr empatar ambos escenarios y estructurar las acciones necesarias para poner en marcha a los comportamientos y subsistemas pertinentes.

El ciclo de inferencia se realiza durante la tarea; el robot puede modificar indefinidamente su plan inicial para completarla. Durante el proceso, el usuario no es capaz de saber el momento en que se inicia el ciclo ni tampoco cuando el robot se percata de que la situación en el entorno es diferente a su concepto general del estado del mundo. El comportamiento emocional y la activación de expresiones faciales pueden informar del progreso de la tarea y cuando las condiciones son consistentes con el estado hipotético del entorno, así mismo, son útiles para comunicar cuando algo se encuentra diferente en el ambiente y en el momento que emerge una situación no esperada. Las expresiones faciales pueden informar acerca de eventos que necesiten de un protocolo de recuperación en el sistema.

A partir de las evaluaciones realizadas durante esta investigación, el enojo puede utilizarse para mostrar que persiste una condición adversa en el entorno, el reinicio del ciclo de inferencia además de inducir a la participación del usuario en caso de que su intervención ayude completar los objetivos.

Finalmente, Golem-III activa las expresiones emocionales en el contexto

de la tarea; las emociones se acoplan objetivamente a favor de la interacción y se reduce la incertidumbre respecto a las acciones del sistema. El ciclo de inferencia es un proceso que complementa la arquitectura cognitiva del robot el cuál se encuentra situado en el entorno de interacción; en este caso, el comportamiento emocional se integra no como una mera ventana del estado interno, sino que su función puede hacer diferencia en el desempeño general del sistema.

## Capítulo 9

# Conclusiones

Las expresiones faciales ocurren siempre durante las actividades sociales. Si la activación es oportuna y destacada, su presencia repentina podría indicar un cambio en la interacción. La activación adecuada de expresiones emocionales dentro de las tareas de un robot de servicio puede mejorar la comunicación e informar acerca de las intenciones y expectativas del sistema. La presente investigación analiza la función de las expresiones faciales universales de mayor legibilidad durante la interacción humano-robot con el fin de establecer su uso y activación durante tareas de servicio a través de modelos de diálogo y un módulo de inferencia implementado en la arquitectura cognitiva del sistema. De esta forma, la activación de emociones en un robot de servicio puede influir el comportamiento del usuario durante la interacción.

El análisis se realizó a partir del diseño y fabricación de una cabeza robótica con elementos estructurales y electromecánicos mínimos capaz de

representar cinco de las seis expresiones faciales universales descritas por Paul Ekman (Felicidad, tristeza, enojo, sorpresa, miedo y disgusto el cuál no fue implementado en el sistema). El prototipo facial desarrollado se presentó a varios sujetos utilizando una encuesta en línea con imágenes del sistema ejecutando las cinco expresiones; la secuencia de las imágenes y respuestas se realizaron de forma aleatoria.

Los resultados obtenidos mostraron que los rostros negativos de tristeza y enojo obtuvieron las mejores puntuaciones de legibilidad; a partir de estos datos se realizaron experimentos enfocados a la interacción humano-robot dentro de escenarios colaborativos. La tarea a realizar fue simple y repetitiva con el fin de facilitar la interacción, diferenciar la reacción del humano a partir del cambio en la activación de las expresiones, evaluar de forma legible la respuesta del usuario y evitar que existan condiciones complejas en la tarea que influyan en el comportamiento humano.

El rostro de enojo obtuvo la puntuación más alta; la expresión facial activada durante la interacción influyó positivamente a favor de la tarea. La emoción activada como retroalimentación después de un fallo programado en el robot incrementa la velocidad de reacción del usuario para ayudar a solventar la situación adversa en la tarea, además regula el comportamiento para ayudar a completar el objetivo general e influye en la atención humana cuando una actividad es confusa y poco clara.

Durante la tarea, el enojo fue activado cuando el robot no fue capaz de completarla al no lograr arrojar un cilindro en un contenedor; éste objeto era inicialmente alcanzado por una persona. La expresión facial tiene un

efecto positivo en la tarea, específicamente, cuando los usuarios observan el rostro siguen las instrucciones adecuadamente. Hipotéticamente, este comportamiento surge a partir de la identificación del enojo en el contexto del fallo e induce al sujeto a colaborar adecuadamente en la tarea; la función social de la emoción informa y orienta al interlocutor humano. Un agente artificial inexpressivo que mantiene una cara neutral, genera situaciones inciertas provocando que el sujeto deduzca la causa del fallo, en este caso, la respuesta del usuario puede afectar de forma positiva o negativa al progreso de la tarea.

Los resultados obtenidos son consistentes con la hipótesis sobre el enojo, en donde se inicia un ciclo de análisis y diagnóstico en el observador; tal ciclo es útil para influir en la atención humana e identificar situaciones poco claras [177]. De acuerdo al Modelo Circumplejo de las Emociones [5, 50–52] y al modelo PA-NA (“*Positive Activation - Negative Activation*”) [53, 54], las emociones son opuestas en diferentes niveles; por ejemplo cuando existe un incremento en el nivel de excitación de las emociones negativas se presenta un cambio de tristeza a enojo. Ambas emociones tienen una condicionante y función diferente [2]. Además, la valencia positiva se muestra en los modelos como contraparte; por ejemplo, tristeza (negativa) a felicidad (positiva). En el caso de enojo, se considera como contraparte positiva a los gestos de contento y relajado, los cuales no son consideradas como emociones universales. Durante la interacción social, las emociones de enojo y tristeza como contraparte de felicidad es poco claro. Sin embargo, durante los experimentos realizados en esta investigación muestran un efecto similar

cuando se utilizan como retroalimentación negativa durante la tarea.

En el experimento principal, se asignó al progreso de la tarea una secuencia de expresiones en el robot. En esta secuencia se activó una transición de rostros que representan el incremento en la intensidad de una emoción. La transición fue utilizada para establecer el contexto de interpretación de un estático; la presencia repentina de un estático durante la activación de transicionales resalta un cambio en el contexto y la valencia de las emociones fue la base para este cambio. Básicamente, se contrastó la transición de positivos con negativos y vice versa. La transición se organizó con expresiones neutrales a positivas (felicidad) o de neutrales a negativas (enojo).

Los resultados mostraron que la presencia repentina de enojo en el contexto de una expresión transicional positiva es efectiva para inducir la atención de sujetos no expertos. Se observó que un amplio número de participantes registraron al menos una causa durante la activación de la emoción. La expresión facial de enojo promueve el análisis y evaluación durante condiciones poco claras que detonan la emoción; el sujeto que la observa indica con mayor precisión la causa de la activación. En contraste, la presencia continua de la expresión es adversa a su legibilidad.

En el contexto del rostro transicional de enojo, la presencia súbita de un positivo fue identificado con mayor precisión que en las pruebas de legibilidad con los estáticos (Capítulo 5). Sin embargo, las expresiones de felicidad y sorpresa fueron menos efectivas para inducir el análisis de acuerdo a la cantidad de causas señaladas durante su activación. El rostro negativo es

estadísticamente diferente de los positivos, en cambio, estos últimos no son diferentes entre sí.

A partir de los estudios en robots de servicio, existen muchas ventajas en la integración de expresiones faciales robóticas, además, las emociones son útiles para orientar la interacción cuando surgen situaciones con intenciones explícitas así como ilegibles; sobre todo cuando se necesita de mensajes directos sin el uso de lenguaje hablado u otro medio. El uso de pocos recursos para comunicar un estado particular del robot puede favorecer a la eficiencia del sistema y a la velocidad para transmitir el mensaje.

Actualmente, estos resultados se han incorporado en el robot de servicio Golem-III [156] (ver figura 5.7) utilizando una cabeza robótica minimalista que incluye una estructura ligera, portátil, fácil de replicar, de bajo mantenimiento y bajo consumo de energía [20]. La activación de emociones se integra en el comportamiento del robot a través de la arquitectura cognitiva y se generan a partir del ciclo de inferencia durante una tarea interactiva.

Los métodos para mostrar expresiones faciales no se limita con el uso de caras robóticas humanoides; muchos otros sistemas utilizan equipos que son capaces de emitir luces, sonidos [194–196] e iconografía cultural para simular un estado emocional [197] (similares a los símbolos utilizados en historietas cómicas) y otros que utilizan luces y sonidos durante tareas de servicio [198,199]. Algunos investigadores implementan el diseño de rostros con elementos físicos que emulan animales a modo de mascotas [34,200,201]. El uso de expresiones faciales humanoides utilizados en este trabajo corresponden a la objetividad de los estudios realizados en otras investigaciones,

al criterio de universalidad y al método de fabricación con el mínimo de elementos para los prototipos faciales.

La activación del comportamiento emocional en agentes artificiales se ha investigado de forma amplia. Los análisis al respecto han inducido al estudio del comportamiento humano que se genera a partir de las expresiones emocionales artificiales, sin embargo, su uso es aún limitado. La base de esta investigación es la integración de las expresiones emocionales de forma objetiva previo al desarrollo de un método para su activación; se tomó como base el entorno de interacción de un robot de servicio en interacción con humanos, la expectativa del usuario, la legibilidad de la expresión y su influencia sobre el comportamiento humano a favor de la tarea. Sin embargo, estos análisis incluyen la activación de emociones durante la ejecución de una actividad cotidiana y se introduce como un comportamiento más en la arquitectura cognitiva. Bajo este criterio, la generación de un comportamiento emocional está sujeta a la interpretación del sistema respecto al entorno, al estado de la tarea y las acciones humanas; en cierto sentido este hecho permite anticipar el comportamiento emocional del robot.

Se espera que un agente emocionalmente inteligente tenga un comportamiento a partir de su concepto del estado del mundo el cuál puede ser diferente a partir de las situaciones cambiantes del entorno social; el humano que interactúa con el sistema puede incluso desconocer si existe una situación diferente a la esperada. Un robot con capacidad para generar un comportamiento emocional a partir de esta divergencia favorece a la interacción e incluso provee de información al usuario en el caso de que este no

se percate de una situación en particular y diferente al estado esperado.

El ciclo de inferencia en Golem-III permite que la interacción fluya de varias maneras. En primer lugar, el sistema da lugar al diagnóstico de la situación seguido de la toma de decisión para después planificar las acciones a realizar. El robot no solo puede informar puntualmente una situación distinta a la esperada, sino que modifica sus acciones para cumplir con un objetivo. En este caso, el comportamiento emocional se acopla al sistema como una ventana a su estado interno, sin embargo, se espera que la expresión informe de manera paralela una condición no esperada.

El ciclo de inferencia es hasta ahora el mejor método para activar emociones en Golem-III ya que se hace explícito el uso de teorías emocionales híbridas como la de la Evaluación y la Dimensional; el agente valora la situación tomando como referente a su estado interno para inferir y reaccionar en consecuencia; la emoción puede cambiar de forma súbita o transicional a partir de la situación. Además, se espera que el comportamiento emocional en el robot genere en el humano una actitud positiva al desarrollo de la tarea de forma similar a lo que se espera en la Teoría de Aproximación Comunicativa. Los modelos computacionales de emoción se exploraron en el capítulo 2.

La aproximación comunicativa se utiliza en los agentes artificiales para advertir al usuario cuando se cometió un acierto o fallo; la teoría es utilizada en sistemas computacionales educativos y con su uso se pretende estimular la atención del usuario. En la teoría se utiliza el manejo de la valencia en las emociones (positiva a negativa y viceversa) para representar el logro o

fracaso en los objetivos. A partir de los hallazgos de esta investigación las expresiones negativas no representan como tal una condición fallida, sino que estimulan al usuario a colaborar y a prestar atención en la situación general de la tarea.

Durante esta investigación se observó de manera no intencional durante entrevistas previas con algunos sujetos, que las expresiones negativas de tristeza y enojo tienen un efecto distinto en las expectativas del usuario. Durante el fallo, los sujetos relacionaron una condición de incapacidad en el sistema cuando se presentó el rostro de tristeza, en el caso de enojo, el usuario asume cierta responsabilidad por la acción desfavorable. Empíricamente la influencia de las expresiones negativas tiene un efecto distinto y es posible sacar provecho a favor de la interacción si se inducen a través de un contexto apropiado, esto se ha estudiado de forma moderada comparando contextos culturales [202].

Durante la interacción social, se espera que exista consistencia con la emoción y la situación a partir de las condiciones que la detonan así como sus funciones [2, 9] (ver cuadro 2.1 en el capítulo 2). La influencia de las expresiones en personas se basa en las emociones generadas por humanos y existe poca información de las emitidas por sistemas artificiales; el efecto de las expresiones emocionales en robots suele configurarse a partir de la interpretación del programador. Como se ha observado hasta ahora, el significado de la expresión corresponde al usuario en relación al entorno, situación, método de ejecución (mecánicos o digitales) y configuración formal del prototipo facial; cómo se observó en el capítulo 5 la modificación

respecto al diseño físico del prototipo facial altera directamente la legibilidad de los rostros, este hecho permite el estudio de patrones formales que favorezcan la legibilidad de expresiones faciales específicas.

La transición de expresiones faciales en los robots de servicio ha sido abordada de forma limitada, sin embargo, se observó que la presencia súbita de un rostro estático destaca de entre los transitorios y tiene mayor significado si se contrasta con la valencia e intensidad; es posible generar entonces el contexto para la activación de expresiones.

Finalmente, el rostro de enojo en un robot tiene una influencia similar al que se genera cuando es expresado por humanos, sin embargo no es claro si el usuario genera una actitud negativa hacia el robot de la misma forma como se hace entre personas [203,204]. La importancia de los presentes hallazgos radica también en el hecho de que las expresiones emocionales negativas implementados en un robot de servicio no necesariamente implican un efecto adverso durante la interacción con los humanos. Por el contrario, las emociones tienen un impacto regulador en una tarea de colaboración e influyen positivamente en el comportamiento del usuario si su activación es oportuna y adecuada a la situación.

Con esta investigación se han desarrollado los siguientes artículos y contribuciones:

- a Reyes, Mauricio., Meza, Iván., Pineda, Luis. Alberto. (2015). The Positive Effect of Negative Feedback in HRI Using a Facial Expression Robot. In International Workshop in Cultural Robotics (pp. 44-54). Springer International Publishing.

- b Reyes, Mauricio., Meza, Iván., Pineda, Luis. Alberto. (2017). Robotics Facial Expression of Anger in Collaborative HRI. International Journal of Advanced Robotic Systems. (Artículo en segunda revisión y aprobado para su publicación.)
- c Pineda, L. A., Rodríguez, A., Fuentes, G., Hernández, N., Reyes, M., Rascón, C., Cruz, R., Vélez, I., Ortega, H. (2017). Opportunistic Inference and Emotions in Service Robots. Journal of Intelligent Fuzzy Systems, IOS Press.
- d Pineda L. A., Rodríguez A., Fuentes G., Rascón C., Reyes M., Ortega H., Hernández N., Cruz R., (2016) The Golem Team, RoboCup@Home 2016. RoboCup German 2016.
- e Pineda L. A., Rodríguez A., Fuentes G., Rascón C., Reyes M., Ortega H., Hernández N., Cruz R., (2017) The Golem Team, RoboCup@Home 2017. RoboCup Nagoya Japan 2017.
- f Participación en taller y conferencia: "The Positive Effect of Negative Feedback in HRI Using a Facial Expression Robot"; dentro del 24vo. Simposio Internacional en Comunicación Humana Interactiva y Robot (Ro-Man) en Kobe Osaka Japón en agosto de 2015.

# Bibliografía

- [1] Plutchik, R. (1984). Emotions: A general psychoevolutionary theory. *Approaches to emotion*, 1984, 197-219.
- [2] Plutchik, R. (1991) *The emotions*, Lanham Maryland: University Press of America.
- [3] Ekman, P. E., Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. Oxford University Press.
- [4] Ortony, A., Clore, G. L., Collins, A. (1990). *The cognitive structure of emotions*. Cambridge university press.
- [5] Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological review*, 110(1), 145.
- [6] Marsella, S., Gratch, J., Petta, P. (2010). Computational models of emotion. *A Blueprint for Affective Computing-A sourcebook and manual*, 11(1), 21-46.

- [7] Marsella, S., Gratch, J. (2014). Computationally modeling human emotion. *Communications of the ACM*, 57(12), 56-67.
- [8] Kriegel, U. (2014). Towards a new feeling theory of emotion. *European Journal of Philosophy*, 22(3), 420-442.
- [9] Breazeal, C. (2004). Function meets style: insights from emotion theory applied to HRI. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 34(2), 187-194.
- [10] Fong, T., Nourbakhsh, I., Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, 42(3), 143-166.
- [11] Midden, C., Ham, J. (2009). Using negative and positive social feedback from a robotic agent to save energy. In *Proceedings of the 4th international conference on persuasive technology (12)*. ACM.
- [12] Kondo, Y., Takemura, K., Takamatsu, J., Ogasawara, T. (2013). A gesture-centric android system for multi-party human-robot interaction. *Journal of Human-Robot Interaction*, 2(1), 133-151.
- [13] Knapp, M. L., Hall, J. A., Horgan, T. G. (2013). *Nonverbal communication in human interaction*. Cengage Learning.
- [14] Bennett C. C., Sabanovic S.(2014) Deriving minimal features for human-like facial expressions in robotic faces., *International Journal of Social Robotics*, 6(3), 367-381.

- [15] Trovato, G., Kishi, T., Endo, N., Hashimoto, K., Takanishi, A. (2012). Development of facial expressions generator for emotion expressive humanoid robot. In *Humanoid Robots (Humanoids), 2012 12th IEEE-RAS International Conference on* (pp. 303-308). IEEE.
- [16] van Breemen, A., Yan, X., Meerbeek, B. (2005). iCat: an animated user-interface robot with personality. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems* (pp. 143-144). ACM.
- [17] Schlosberg, H. (1954). Three dimensions of emotion. *Psychological review*, 61(2), 81.
- [18] Goodrich, M. A., Schultz, A. C. (2007). Human-robot interaction: a survey. *Foundations and trends in human-computer interaction*, 1(3), 203-275.
- [19] Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human-robot interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1480), 679-704.
- [20] Pineda, L. A., Rodríguez, A., Fuentes, G., Rascon, C., Meza, I. V. (2015) Concept and functional structure of a service robot. *International Journal of Advanced Robotic Systems*, 12(2), 6.
- [21] Pineda, L. A., Salinas, L., Meza, I. V., Rascon, C., Fuentes, G. (2013). Sitlog: a programming language for service robot tasks. In-

- ternational Journal of Advanced Robotic Systems, 10(10), 358. DOI: <https://doi.org/10.5772/56906>
- [22] Pineda, L. A. (2008). Specification and interpretation of multimodal dialogue models for human-robot interaction. *Artificial Intelligence for Humans: Service Robots and Social Modeling*, 33-50.
- [23] Maria Andries, A. (2011) Positive and Negative Emotions within the Organizational Context. *Global Journal of Human-Social Science Research*, 11(9).
- [24] Solomon, R. C., Stone, L. D. (2002) On “positive” and “negative” emotions. *Journal for the Theory of Social Behaviour*, 32(4), 417-435.
- [25] Pratto, F., John, O. P. (1991) Automatic Vigilance: The Attention-Grabbing Power of negative Social Information. *Journal of Personality and Social Psychology*, 161 (3), 380-391.
- [26] Eastwood, J. D., Smilek, D., Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Attention, Perception, & Psychophysics*, 65(3), 352-358.
- [27] Boucher, J. D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S., Ventre-Dominey, J. (2012). I reach faster when I see you look: gaze effects in human-human and human-robot face-to-face cooperation. *Frontiers in neurorobotics*, 6.
- [28] Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., Berlin, M. (2005, August). Effects of nonverbal communication on efficiency and

- robustness in human-robot teamwork. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on* (pp. 708-713). IEEE.
- [29] Sabelli, A. M., Kanda, T. (2016). Robovie as a mascot: a qualitative study for long-term presence of robots in a shopping mall. *International Journal of Social Robotics*, 8(2), 211-221.
- [30] Thrun, S., Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., Schulz, D. (1999) MINERVA: A second-generation museum tour-guide robot. In *Robotics and automation, 1999. Proceedings. 1999 IEEE international conference on*, 3.
- [31] Nieuwenhuisen, M., Behnke, S. (2013). Human-like interaction skills for the mobile communication robot robotinho. *International Journal of Social Robotics*, 5(4), 549-561.
- [32] Sell, A., Tooby, J., Cosmides, L. (2009) Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences*, 106(35), 15073-15078.
- [33] Sell, A. (2005) *Regulating welfare tradeoff ratios: three tests of an evolutionary-computational model of human anger* (Doctoral dissertation, University of California Santa Barbara).
- [34] Breazeal, C. (2003) Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1), 119-155.

- [35] Norman, D. A., Ortony, A., Russell, D. M. (2003). Affect and machine design: Lessons for the development of autonomous machines. *IBM Systems Journal*, 42(1), 38-44.
- [36] Ekman, P. (1973). Cross-cultural studies of facial expression. *Darwin and facial expression: A century of research in review*, 169-222.
- [37] Ekman P. (1999) Basic emotions, *Handbook of Cognition and Emotion*, T. Dalgleish and M.
- [38] Russell, J. A. (1980) A circumplex model of affect. *J. Personality and Social Psychology*, 39, 1161-78.
- [39] Ortony, A., Turner, T. J. (1990). What's basic about basic emotions?. *Psychological review*, 97(3), 315.
- [40] Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, 6(3-4), 169-200.
- [41] Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. *Cognition and emotion*, 1(2), 115-143.
- [42] Frijda, N. H., Swagerman, J. (1987). Can computers feel? Theory and design of an emotional system. *Cognition and emotion*, 1(3), 235-257.
- [43] Frijda, N. H. (1988). The laws of emotion. *American psychologist*, 43(5), 349.

- [44] Velásquez, J. (1998, August). A computational framework for emotion-based control. In Proceedings of the Workshop on Grounding Emotions in Adaptive Systems; International Conference on SAB (pp. 62-67).
- [45] Clore, G. L., Ortony, A. (2000). Cognition in emotion: Always, sometimes, or never. *Cognitive neuroscience of emotion*, 24-61.
- [46] Ekman, P., Friesen, W. V. (2003) *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk.
- [47] Spindler, O., Fadrus, T. (2009). Grimace project documentation. Vienna University of Technology. Recuperado de <http://www.grimace-project.net/>
- [48] Cohn, J. F. (2010). Advances in behavioral science using automated facial image analysis and synthesis [social sciences]. *IEEE Signal processing magazine*, 27(6), 128-133.
- [49] Reyes, M., Meza, I., Pineda, L. A. (2015). The positive effect of negative feedback in HRI using a facial expression robot. In *International Workshop in Cultural Robotics* (pp. 44-54). Springer International Publishing.
- [50] Russell, J. A., Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of personality and social psychology*, 76(5), 805.
- [51] Russell, J. A., Carroll, J. M. (1999) On the bipolarity of positive and negative affect. *Psychological bulletin*, 125(1), 3.

- [52] Posner, J., Russell, J. A., Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(3), 715-734.
- [53] Watson, D., Tellegen, A. (1985) Toward a consensual structure of mood. *Psychological bulletin*, 98(2), 219.
- [54] Watson, D., Wiese, D., Vaidya, J., Tellegen, A. (1999) The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of personality and social psychology*, 76(5), 820.
- [55] Larsen, J. T., McGraw, A. P. (2011) Further evidence for mixed emotions. *Journal of personality and social psychology*, 100(6), 1095.
- [56] Eastwood, J. D., Smilek, D., Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Attention, Perception, & Psychophysics*, 63(6), 1004-1013.
- [57] Midden, C., Ham, J. (2014) The Power of Negative Feedback from an Artificial Agent to Promote Energy Saving Behavior. In *International Conference of Design, User Experience, and Usability*, Springer International Publishing, 328-338.
- [58] Lazarus, R. S. (1991). Cognition and motivation in emotion. *American psychologist*, 46(4), 352.

- [59] Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press on Demand.
- [60] Scherer, K. R., Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal?. *Emotion*, 7(1), 113.
- [61] Mehrabian, A., Russell, J. A. (1974). *An approach to environmental psychology*. the MIT Press.
- [62] Gratch, J. (2008). True emotion vs. social intentions in nonverbal communication: Towards a synthesis for embodied conversational agents. *Modeling communication with robots and virtual humans*, 181-197.
- [63] Simon, H. A. (1967). Motivational and emotional controls of cognition. *Psychological Review* 74: 29-39.
- [64] Sloman, A., Croucher, M. (1981). Why robots will have emotions. *International Joint Conference on Artificial Intelligence*, Vancouver, Canada.
- [65] Frank, R. (1988). *Passions with reason: the strategic role of the emotions*. New York, NY, W. W. Norton.
- [66] Scheutz, M., Sloman, A. (2001). *Affect and agent control: experiments with simple affective states*. IAT, World Scientific Publisher.

- [67] Doyle, J. (2006). *Extending Mechanics to Minds: The Mechanical Foundations of Psychology and Economics*. London, UK, Cambridge University Press.
- [68] Marsella, S., Gratch, J. (2006). EMA: A computational model of appraisal dynamics. na.
- [69] Gratch, J., Marsella, S. (2004). A domain independent framework for modeling emotion. *Journal of Cognitive Systems Research* 5(4): 269-306.
- [70] Gratch, J., Marsella, S. (2004). Evaluating the modeling and use of emotion in virtual humans. 3rd International Joint Conference on Autonomous Agents and Multiagent Systems, New York.
- [71] Marsella, S., Gratch, J., Marsella, S., Rickel, J. (2004). Expressive behaviors for virtual worlds. In *Life-like characters* (pp. 317-360). Springer Berlin Heidelberg.
- [72] Gebhard, P. (2005). ALMA - A Layered Model of Affect. Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, Utrecht.
- [73] Marinier, R. P. (2008). *A Computational Unification of Cognitive Control, Emotion, and Learning*. Computer Science. Ann Arbor, MI, University of Michigan. PhD.
- [74] Elliott, C. (1992). *The affective reasoner: A process model of emotions in a multi-agent system*. Northwestern, IL, Northwestern University Institute for the Learning Sciences.

- [75] Becker-Asano, C., Wachsmuth, I. (2008). Affect Simulation with Primary and Secondary Emotions. 8th International Conference on Intelligent Virtual Agents, Tokyo, Springer.
- [76] Becker-Asano, C. (2008). WASABI: Affect simulation for agents with believable interactivity. Bielefeld, University of Bielefeld. PhD.
- [77] Reisenzein, R. (2009). "Emotions as Metarepresentational States of Mind: Naturalizing the Belief-Desire Theory of Emotion." *Journal of Cognitive Systems Research* 10(1).
- [78] Neal Reilly, W. S. (2006). Modeling what happens between emotional antecedents and emotional consequents. Eighteenth European Meeting on Cybernetics and Systems Research. Vienna, Austria, Austrian Society for Cybernetic Studies: 607-612.
- [79] Bui, T. D. (2004). Creating emotions and facial expressions for embodied agents. Department of Electrical Engineering, Mathematics and Computer Science. Enschede, University of Twente. PhD
- [80] Gmytrasiewicz, P. Lisetti, C. (2000). Using Decision Theory to Formalize Emotions for Multi-Agent Systems. Second ICMAS-2000 Workshop on Game Theoretic and Decision Theoretic Agents, Boston.
- [81] Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1535), 3453-3458.

- [82] Blow, M., Dautenhahn, K., Appleby, A., Nehaniv, CL., Lee D.C. (2006) Perception of robot smiles and dimensions for human-robot interaction design. 15th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN), pp 469–474.
- [83] Mori, M. (1970) Bukimi no tani “The uncanny valley”. *Energy* 7(4):33–35.
- [84] MacDorman. K.F., Green, R.D., Ho, C-C., Koch, C.T. (2009) Too real for comfort? Uncanny responses to computer generated faces. *Computer Human Behavior* 25(3):695–710.
- [85] Mathur, M. B., Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition*, 146, 22-32.
- [86] Burleigh, T. J., Schoenherr, J. R., Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*, 29(3), 759-771.
- [87] DiSalvo CF, Gemperle F, Forlizzi J, Kiesler S (2002) All robots are not created equal: the design and perception of humanoid robot heads. *Proceedings of the 4th ACM Conference on Designing Interactive Systems*, pp 321–326.
- [88] Seib V., Giesen J., Grüntjens D., Paulus D. (2013) Enhancing human-robot interaction by a robot face with facial expressions and synchronized

- lip movements., Communication Papers Proceedings: 21st International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision in co-operation with EUROGRAPHICS Association,
- [89] Moubayed S. A., Skantze G., Beskow J. (2013) The furhat back-projected humanoid head–lip reading, gaze and multi-party interaction.» International Journal of Humanoid Robotics, 10(01)
- [90] Itoh, K., Miwa, H., Matsumoto, M., Zecca, M., Takanobu, H., Roccella, S., Takanishi, A. (2004) Various emotional expressions with emotion expression humanoid robot WE-4RII. In Robotics and Automation, 2004. TExCRA'04. First IEEE Technical Exhibition Based Conference on (35-36). IEEE.
- [91] Kishi, T., Kojima, T., Endo, N., Destephe, M., Otani, T., Jamone, L., Takanishi, A. (2013) Impression survey of the emotion expression humanoid robot with mental model based dynamic emotions. In Robotics and Automation (ICRA), 2013 IEEE International Conference on (1663-1668). IEEE.
- [92] Hegel F., Eyssel F., Wrede B. (2010) The social robot “flobi”: Key concepts of industrial design. In 19th International Symposium in Robot and Human Interactive Communication. IEEE., 107-112.
- [93] Bazo, D., Vaidyanathan, R., Lentz, A., Melhuish, C. (2010, October). Design and testing of a hybrid expressive face for a humanoid robot. In

Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on (pp. 5317-5322). IEEE.

- [94] Breazeal, C. (2003). Toward sociable robots. *Robotics and autonomous systems*, 42(3), 167-175.
- [95] Becker-Asano, C., Ishiguro, H. (2011). Evaluating facial displays of emotion for the android robot Geminoid F. In *Affective Computational Intelligence (WACI), 2011 IEEE Workshop on* (pp. 1-8). IEEE.
- [96] Berns, K., Hirth, J. (2006) Control of facial expressions of the humanoid robot head ROMAN. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp 3119–3124.
- [97] Miwa H., Umetsu T., Takanishi A., Takanobu H.(2001) Robot personality based on the equations of emotion defined in the 3D mental space, In *Robotics and Automation, Proceedings 2001 ICRA. IEEE International Conference*, 3, 2602-260.
- [98] Miwa, H., Itoh, K., Matsumoto, M., Zecca, M., Takanobu, H., Rocella, S., Takanishi, A. (2004, September). Effective emotional expressions with expression humanoid robot we-4rii: integration of humanoid robot hand rch-1. In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on* (Vol. 3, pp. 2203-2208). IEEE.
- [99] Kishi, T., Otani, T., Endo, N., Kryczka, P., Hashimoto, K., Nakata, K., Takanishi, A. (2012, October). Development of expressive robotic head

- for bipedal humanoid robot. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (pp. 4584-4589). IEEE.
- [100] Trovato, G., Kishi, T., Endo, N., Hashimoto, K., Takanishi, A. (2012). A cross-cultural study on generation of culture dependent facial expressions of humanoid social robot. *Social Robotics*, 35-44.
- [101] Nieuwenhuisen, M., Behnke, S. (2013). Human-like interaction skills for the mobile communication robot robotinho. *International Journal of Social Robotics*, 5(4), 549-561.
- [102] Allman, T. (2009). *The Nexi Robot*. Norwood House Press.
- [103] Ferland, F., Létourneau, D., Aumont, A., Frémy, J., Legault, M. A., Lauria, M., Michaud, F. (2012). Natural interaction design of a humanoid robot. *Journal of Human-Robot Interaction*, 1(2), 118-134.
- [104] Wainer, J., Dautenhahn, K., Robins, B., Amirabdollahian, F. (2014). A pilot study with a novel setup for collaborative play of the humanoid robot KASPAR with children with autism. *International journal of social robotics*, 6(1), 45-65.
- [105] Ruiz-del-Solar, J., Correa, M., Verschae, R., Bernuy, F., Loncomilla, P., Mascaró, M., Smith, F. (2012) Bender: A General-Purpose Social Robot with Human-Robot Interaction Capabilities. *Journal of Human-Robot Interaction*, 1(2), 54-75.
- [106] Seeger, M., Ultra-Large-Sites, S. (2009). Key-Value stores: a practical overview. *Computer Science and Media*, Stuttgart.

- [107] Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., Akamatsu, S. (2013) Dynamic properties influence the perception of facial expressions. *Perception*, 42(11), 1266-1278.
- [108] Ambadar, Z., Schooler, J. W., Cohn, J. F. (2005) Deciphering the enigmatic face the importance of facial dynamics in interpreting subtle facial expressions. *Psychological science*, 16(5), 403-410.
- [109] Ekman, P., Rosenberg, E. L. (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, USA.
- [110] Dautenhahn, K. (2002). Design spaces and niche spaces of believable social robots. In *Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on* (pp. 192-197). IEEE.
- [111] Dautenhahn, K. (2003). Roles and functions of robots in human society: implications from research in autism therapy. *Robotica*, 21(4), 443-452.
- [112] Pollack, M. E., Brown, L., Colbry, D., Orosz, C., Peintner, B., Ramakrishnan, S., Thrun, S. (2002, August). Pearl: A mobile robotic assistant for the elderly. In *AAAI workshop on automation as eldercare* (Vol. 2002, pp. 85-91).
- [113] Pineau, J., Montemerlo, M., Pollack, M., Roy, N., Thrun, S. (2003). Towards robotic assistants in nursing homes: Challenges and results. *Robotics and autonomous systems*, 42(3), 271-281.

- [114] Fong, T., Thorpe, C., Baur, C. (2001). Collaboration, dialogue and human-robot interaction, 10th international symposium of robotics research (lorne, victoria, australia). In Proceedings of the 10th International Symposium of Robotics Research.
- [115] Buhmann, J., Burgard, W., Cremers, A. B., Fox, D., Hofmann, T., Schneider, F. E., Thrun, S. (1995). The mobile robot Rhino. *Ai Magazine*, 16(2), 31.
- [116] Schulte, J., Rosenberg, C., Thrun, S. (1999). Spontaneous, short-term interaction with mobile robots. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on* (Vol. 1, pp. 658-663). IEEE.
- [117] Willeke, T., Kunz, C., Nourbakhsh, I. R. (2001). The History of the Mobot Museum Robot Series: An Evolutionary Study. In *FLAIRS Conference* (pp. 514-518).
- [118] Bruce, A., Nourbakhsh, I., Simmons, R. (2002). The role of expressiveness and attention in human-robot interaction. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on* (Vol. 4, pp. 4138-4142). IEEE.
- [119] Siegwart, R., Arras, K. O., Bouabdallah, S., Burnier, D., Froidevaux, G., Greppin, X., Philippsen, R. (2003). Robox at Expo. 02: A large-scale installation of personal robots. *Robotics and Autonomous Systems*, 42(3), 203-222.

- [120] Ishiguro, H., Ono, T., Imai, M., Maeda, T., Kanda, T., Nakatsu, R. (2001). Robovie: an interactive humanoid robot. *Industrial robot: An international journal*, 28(6), 498-504.
- [121] Shiomi, M., Kanda, T., Ishiguro, H., Hagita, N. (2006). Interactive humanoid robots for a science museum. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction* (pp. 305-312). ACM.
- [122] Yamaoka, F., Kanda, T., Ishiguro, H., Hagita, N. (2010). A model of proximity control for information-presenting robots. *IEEE Transactions on Robotics*, 26(1), 187-195.
- [123] Bennewitz, M., Faber, F., Joho, D., Schreiber, M., Behnke, S. (2005). Towards a humanoid museum guide robot that interacts with multiple persons. In *Humanoid Robots, 2005 5th IEEE-RAS International Conference on* (pp. 418-423). IEEE.
- [124] Faber, F., Bennewitz, M., Eppner, C., Gorog, A., Gonsior, C., Joho, D., Behnke, S. (2009). The humanoid museum tour guide Robotinho. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on* (pp. 891-896). IEEE.
- [125] Huang, C. M., Iio, T., Satake, S., Kanda, T. (2014). Modeling and Controlling Friendliness for An Interactive Museum Robot. In *Robotics: Science and Systems*.

- [126] Díaz, M., Paillacho, D., Angulo, C., Torres, O., González, J., Albo-Canals, J. (2014). A week-long study on Robot-Visitors Spatial Relationships during Guidance in a Sciences Museum. In Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction (pp. 152-153). ACM.
- [127] Díaz-Boladeras, M., Paillacho, D., Angulo, C., Torres, O., González-Diéguéz, J., Albo-Canals, J. (2015). Evaluating group-robot interaction in crowded public spaces: A week-long exploratory study in the wild with a humanoid robot guiding visitors through a science museum. *International Journal of Humanoid Robotics*, 12(04), 1550022.
- [128] Ramón-Vigo, R., Pérez-Higueras, N., Merino, L., Caballero, F. (2014). Transferencia de comportamientos de navegación humana a un planificador local de robots. XXXV Jornadas de Automática.
- [129] Karreman, D., Ludden, G., Evers, V. (2015). Visiting cultural heritage with a tour guide robot: a user evaluation study in-the-wild. In *International Conference on Social Robotics* (pp. 317-326). Springer, Cham.
- [130] Das, D., Rashed, M. G., Kobayashi, Y., Kuno, Y. (2014, March). Recognizing gaze pattern for human robot interaction. In Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction (pp. 142-143). ACM.
- [131] Rashed, M. G., Suzuki, R., Lam, A., Kobayashi, Y., Kuno, Y. (2015). Toward museum guide robots proactively initiating interaction with hu-

- mans. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts (pp. 1-2). ACM.
- [132] Duffy, B. R., Rooney, C., O'Hare, G. M., O'Donoghue, R. (1999). What is a Social Robot?. In 10th Irish Conference on Artificial Intelligence and Cognitive Science, University College Cork, Ireland, 1-3.
- [133] Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and autonomous systems*, 42(3), 177-190.
- [134] Breazeal, C., Scassellati, B. (1999). A context-dependent attention system for a social robot. *rn*, 255, 3.
- [135] Atkeson, C. G., Hale, J. G., Pollick, F., Riley, M., Kotosaka, S., Schaul, S., Kawato, E. (2000). Using humanoid robots to study human behavior. *IEEE Intelligent Systems and their applications*, 15(4), 46-56.
- [136] Restivo, S. (2002). Romancing the robots: Social robots and society. In Proceedings of the Robots as Partners: An Exploration of Social Robots Workshop, International Conference on Intelligent Robots and Systems (IROS 2002), Lausanne, Switzerland.
- [137] Breazeal, C. L. (2004). Designing sociable robots. MIT press.
- [138] Wada, K., Shibata, T., Saito, T., Tanie, K. (2002). Analysis of factors that bring mental effects to elderly people in robot assisted activity. In *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on (Vol. 2, pp. 1152-1157)*. IEEE.

- [139] Dautenhahn, K., Werry, I. (2000). Issues of robot-human interaction dynamics in the rehabilitation of children with autism. Proc. From animals to animats, 6, 519-528.
- [140] Sabanovic, S., Michalowski, M. P., Simmons, R. (2006). Robots in the wild: Observing human-robot social interaction outside the lab. In Advanced Motion Control, 2006. 9th IEEE International Workshop on (pp. 596-601). IEEE.
- [141] Hegel, F., Krach, S., Kircher, T., Wrede, B., Sagerer, G. (2008). Understanding social robots: A user study on anthropomorphism. In Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on (pp. 574-579). IEEE.
- [142] Lim, G. H., Suh, I. H., Suh, H. (2011). Ontology-based unified robot knowledge for service robots in indoor environments. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 41(3), 492-509.
- [143] Kawamura, K., Pack, R. T., Iskarous, M. (1995). Design philosophy for service robots. In Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century., IEEE International Conference on (Vol. 4, pp. 3736-3741). IEEE.
- [144] Hertzberg, J., Saffiotti, A., (2008) Using semantic knowledge in robotics, Robotics and Autonomous Systems, vol. 56, no. 11, pp. 875–877.

- [145] Severinson-Eklundh, K., Green, A., Hüttenrauch, H. (2003). Social and collaborative aspects of interaction with a service robot. *Robotics and Autonomous systems*, 42(3), 223-234.
- [146] Russell, S. J., Norvig, P. (2002). *Artificial intelligence: a modern approach* (International Edition).
- [147] Suh, I. H., Lim, G. H., Hwang, W., Suh, H., Choi, J. H., Park, Y. T. (2007). Ontology-based multi-layered robot knowledge framework (OMRKF) for robot intelligence. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on* (pp. 429-436). IEEE.
- [148] Wisspeintner, T., Van Der Zant, T., Iocchi, L., Schiffer, S. (2009). RoboCup@ Home: Scientific competition and benchmarking for domestic service robots. *Interaction Studies*, 10(3), 392-426.
- [149] Stückler, J., Badami, I., Droeschel, D., Gräve, K., Holz, D., McElhonne, M., Behnke, S. (2013). Nimbro@ home: Winning team of the robo-cup@ home competition 2012. In *RoboCup 2012: Robot Soccer World Cup XVI* (pp. 94-105). Springer, Berlin, Heidelberg.
- [150] van der Burgh, M. F. B., Lunenburg, J. J. M., Appeldoorn, R. P. W., Wijnands, R. W. J., Clephas, T. T. G., Baeten, M. J. J., van de Molengraft, M. J. G. (2017). Tech united eindhoven@ home 2017 team description paper. University of Technology Eindhoven.

- [151] Savage, J., Negrete, M., Matamoros, M., Cruz, J., Contreras, L., Pacheco, A., Márquez, J. (2013). Pumas@ Home 2013 Team Description Paper.
- [152] Memmesheimer, R., Seib, V., Heuer, G., Schmidt, P., Thies, D., Mykhalchyshyna, I., Schütz, R. RoboCup 2016-homer@ UniKoblenz (Germany).
- [153] Hori, S., Ishida, Y., Kiyama, Y., Tanaka, Y., Kuroda, Y., Hisano, M. Hashimoto, K. Hibikino-Musashi@ Home SPL 2017 Team Description Paper.
- [154] Pavez, M., Munoz, R., Gomez, C., Bano, D., Pavez, L., Astorga, J., Ruiz-del-Solar, J. UChile Peppers 2017 Team Description Paper.
- [155] Pineda, L. A., Rascon, C., Fuentes, G. The Golem Team, RoboCup@ Home 2014, Team Description Paper.
- [156] Pineda L. A., Rodriguez A., Fuentes G., Rascon C., Reyes M., Ortega H., Hernandez N., Cruz R., (2016) The Golem Team, RoboCup@ Home 2016.
- [157] Pineda, L. A., Rascon, C., Fuentes, G., (2017) The Golem Team, RoboCup@ Home 2017, Team Description Paper.
- [158] Pineda, L., Meza, I., Aviles, H., Gershenson, C., Rascon, C., Alvarado, M., Salinas, L. (2011). IOCA: interactionoriented cognitive architecture. *Research in Computer Science*, 54, 273-284.

- [159] Allen, J. F., Byron, D. K., Dzikovska, M., Ferguson, G., Galescu, L., Stent, A. (2001). Toward conversational human-computer interaction. *AI magazine*, 22(4), 27.
- [160] D'Ausilio, A. (2012). Arduino: A low-cost multipurpose lab equipment. *Behavior research methods*, 44(2), 305-313.
- [161] Badamasi, Y. A. (2014). The working principle of an Arduino. In *Electronics, Computer and Computation (ICECCO), 2014 11th International Conference on* (pp. 1-4). IEEE.
- [162] Pham, D. T., Gault, R. S. (1998). A comparison of rapid prototyping technologies. *International Journal of machine tools and manufacture*, 38(10), 1257-1287.
- [163] Mueller, S., Kruck, B., Baudisch, P. (2013). LaserOrigami: laser-cutting 3D objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2585-2592). ACM.
- [164] Gershenfeld, N. (2008). *Fab: the coming revolution on your desktop—from personal computers to personal fabrication*. Basic Books.
- [165] Gibson, I., Rosen, D., Stucker, B. (2014). *Additive manufacturing technologies: 3D printing, rapid prototyping, and direct digital manufacturing*. Springer.
- [166] Choudhury, I. A., Shirley, S. (2010). Laser cutting of polymeric materials: an experimental investigation. *Optics and Laser Technology*, 42(3), 503-508.

- [167] Sass, L., Oxman, R. (2006). Materializing design: the implications of rapid prototyping in digital design. *Design Studies*, 27(3), 325-355.
- [168] Yeong, W. Y., Chua, C. K., Leong, K. F., Chandrasekaran, M. (2004). Rapid prototyping in tissue engineering: challenges and potential. *TRENDS in Biotechnology*, 22(12), 643-652.
- [169] Roubroeks, M., Ham, J., Midden, C. (2011) When artificial social agents try to persuade people: The role of social agency on the occurrence of psychological reactance. *International Journal of Social Robotics*, 3(2), 155-165.
- [170] Midden, C., Ham, J. (2012). The illusion of agency: the influence of the agency of an artificial agent on its persuasive power. *Persuasive technology. Design for health and safety*, 90-99.
- [171] Ham, J., Cuijpers, R. H., Cabibihan, J. J. (2015). Combining robotic persuasive strategies: the persuasive power of a storytelling robot that uses gazing and gestures. *International Journal of Social Robotics*, 7(4), 479-487.
- [172] Hamari, J., Koivisto, J., Pakkanen, T. (2014). Do persuasive technologies persuade?-a review of empirical studies. In *International Conference on Persuasive Technology* (pp. 118-136). Springer, Cham
- [173] Scheutz, M., Schermerhorn, P., Kramer, J. (2006). The utility of affect expression in natural language interactions in joint human-robot tasks.

In HRI 2006 Proceedings of the 1st ACM SIGCHI/SIGART, Salt Lake City, Utah, USA, pp. 226–233

- [174] Samer, A., Gabriel, S.(2012) Perception of gaze direction for situated interaction. In 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction. DiVa, Santa Monica
- [175] Haddadi, A., Crof, E., Gleeson, B., MacLean, K., Alcazar, J.(2013). Analysis of task-based gestures in human-robot interaction. In IEEE International Conference on Robotics and Automation (ICRA) Karlsruhe, Germany, pp. 2138–2144. IEEE
- [176] Nomura, T., Kanda, T., Suzuki, T. (2006). Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *AI Soc.* 20(2), 138–150
- [177] Slovic, P., Finucane, M. L., Peters, E., MacGregor, D. G. (2004) Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk analysis*, 24(2), 311-322
- [178] Gallagher, S., Varga, S. (2014) Social constraints on the direct perception of emotions and intentions. *Topoi*, 33(1), 185-199
- [179] Johnson, S. C. (2000) The recognition of mentalistic agents in infancy. *Trends in cognitive sciences*, 4(1), 22-28.
- [180] Lee, H. S., Park, J. W., Chung, M. J. (2007) A linear affect–expression space model and control points for mascot-type facial robots. *IEEE Transactions on Robotics*, 23(5), 863-873.

- [181] Tooby, J., Cosmides, L., Sell, A., Lieberman, D., Sznycer, D. (2008) Internal regulatory variables and the design of human motivation: A computational and evolutionary approach. *Handbook of approach and avoidance motivation*, 15, 251.
- [182] Ekman, P., Davidson, R. J., Friesen, W. V. (1990) The Duchenne smile: Emotional expression and brain physiology: II. *Journal of personality and social psychology*, 58(2), 342.
- [183] Gunnery, S. D., Ruben, M. A. (2016) Perceptions of Duchenne and non-Duchenne smiles: A meta-analysis. *Cognition and Emotion*, 30(3), 501-515.
- [184] Da Pos, O., Green-Armytage, P. (2012) Facial expressions, colours and basic emotions. *JAIC-Journal of the International Colour Association*, 1.
- [185] Barrick, C. B., Taylor, D., Correa, E. I. (2002) Color sensitivity and mood disorders: biology or metaphor?. *Journal of affective disorders*, 68(1), 67-71.
- [186] Kim, M. G., Lee, H. S., Park, J. W., Jo, S. H., Chung, M. J. (2008) Determining color and blinking to support facial expression of a robot for conveying emotional intensity, In *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication*, 219-224.

- [187] Pineda, L. A., Rodríguez, A., Fuentes, G., Rascón, C., Meza, I. (2017). A light non-monotonic knowledge-base for service robots. *Intelligent Service Robotics*, 10(3), 159-171.
- [188] Pineda, L. A., Rodríguez, A., Fuentes, G., Hernández, N., Reyes, M., Rascón, C., Cruz, R., Vélez, I., Ortega, H. (2018). Opportunistic inference and emotion in service robots. *Journal of Intelligent & Fuzzy Systems*, 34(5), 3301-3311
- [189] Brenner, M. (2007). Situation-aware interpretation, planning and execution of user commands by autonomous robots. In *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, 540-545.
- [190] Schiffer, S., Ferrein, A., Lakemeyer, G. (2012). Caesar: an intelligent domestic service robot. *Intelligent Service Robotics*, 5(4), 259-273.
- [191] Beetz, M., Mösenlechner, L., Tenorth, M. (2010). CRAM—A Cognitive Robot Abstract Machine for everyday manipulation in human environments. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, 1012-1017.
- [192] Lemaignan, S., Warnier, M., Sisbot, E. A., Clodic, A., Alami, R. (2017). Artificial cognition for social human–robot interaction: An implementation. *Artificial Intelligence*, 247, 45-69.

- [193] Agostini, A., Torras, C., Wörgötter, F. (2017). Efficient interactive decision-making framework for robotic applications. *Artificial Intelligence*, 247, 187-212.
- [194] Kim, M. G., Lee, H. S., Park, J. W., Jo, S. H., Chung, M. J. (2008). Determining color and blinking to support facial expression of a robot for conveying emotional intensity. In *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on* (pp. 219-224). IEEE.
- [195] Kim, H. R., Kwon, D. S. (2010). Computational model of emotion generation for human-robot interaction based on the cognitive appraisal theory. *Journal of Intelligent & Robotic Systems*, 60(2), 263-283.
- [196] Park, J. C., Song, H., Koo, S., Kim, Y. M., Kwon, D. S. (2010). Robot's behavior expressions according to the sentence types and emotions with modification by personality. In *Advanced Robotics and its Social Impacts (ARSO), 2010 IEEE Workshop on* (pp. 105-110). IEEE.
- [197] Trovato, G., Kishi, T., Endo, N., Hashimoto, K., Takanishi, A. (2012). A cross-cultural study on generation of culture dependent facial expressions of humanoid social robot. *Social Robotics*, 35-44.
- [198] Baraka, K., Rosenthal, S., Veloso, M. (2016). Enhancing human understanding of a mobile robot's state and actions using expressive lights. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on* (pp. 652-657). IEEE.

- [199] Baraka, K., Paiva, A., Veloso, M. (2016). Expressive lights for revealing mobile service robot state. In *Robot 2015: Second Iberian Robotics Conference* (pp. 107-119). Springer International Publishing.
- [200] Sosnowski, S., Bittermann, A., Kuhlentz, K., Buss, M. (2006). Design and evaluation of emotion-display EDDIE. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on* (pp. 3113-3118). IEEE.
- [201] Saldien, J., Goris, K., Vanderborght, B., Vanderfaeillie, J., Lefeber, D. (2010). Expressing emotions with the social robot Probo. *International Journal of Social Robotics*, 2(4), 377-389.
- [202] Bennett, C. C. (2015). The effects of culture and context on perceptions of robotic facial expressions. *Interaction Studies*, 16(2), 272-302.
- [203] Harmon-Jones, E., Sigelman, J. (2001). State anger and prefrontal brain activity: evidence that insult-related relative left-prefrontal activation is associated with experienced anger and aggression. *Journal of personality and social psychology*, 80(5), 797.
- [204] Tiedens, L. Z. (2001). Anger and advancement versus sadness and subjugation: the effect of negative emotion expressions on social status conferral. *Journal of personality and social psychology*, 80(1), 86.

## Apéndice A

# Conceptos y desarrollo de los prototipos faciales

El diseño y desarrollo del prototipo facial de Golem-III fue sujeto a diversas modificaciones que respondieron a las demandas del proyecto. El concepto general de los robots desarrollados en el Proyecto Golem integra los equipos funcionales con la configuración formal simple utilizando algunos elementos estéticos representativos de la cultura prehispánica y de la Universidad Nacional Autónoma de México.

El concepto de los robots Golem pretende mostrar la integración de tecnología con diseño industrial; con el fin de establecer las bases para la fabricación industrial y registro de patentes con los elementos desarrollados. En la figura A.1 se muestra la evolución de los sistemas Golem (Golem-II, Golem-II+ y Golem-III).

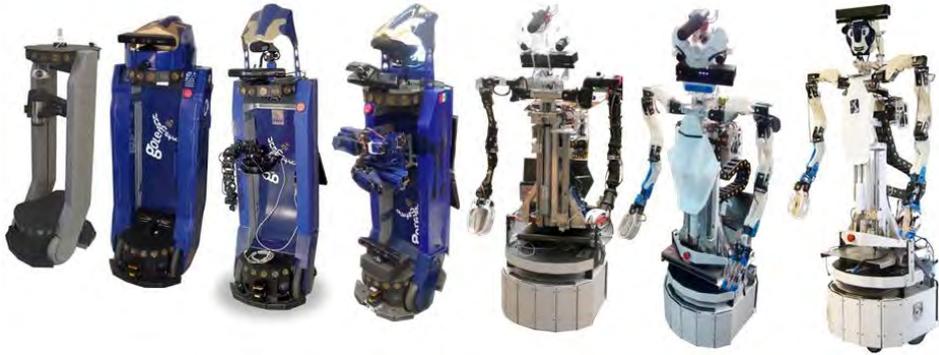


Figura A.1: Evolución de los robots de servicio Golem-II y Golem-III.

A lo largo de su desarrollo se intentó conservar la configuración formal general; se han mantenido elementos similares en el diseño de cabeza y rostro. La memoria gráfica que se presenta a continuación describe el trabajo realizado para el prototipo facial de Golem-III desde los primeros conceptos hasta el diseño final capaz de ejecutar las expresiones faciales universales. El periodo abarca desde inicios de 2014 a la primera mitad de 2017. Las figuras muestran los bocetos iniciales, ilustraciones a detalle, desarrollo de planos de trabajo, modelado digital tridimensional, modelos físicos de presentación elaborados en sistemas para la fabricación de prototipos rápidos y prototipos funcionales.

### A.1. Memoria gráfica descriptiva.

La figura A.2 muestra el diseño inicial para el rostro de Golem-III; boceto generado en el programa “Sketchbook Pro V.3” mostrando las ex-



Figura A.2: Concepto inicial Golem-III y las expresiones básicas universales

presiones universales básicas.

En las figuras A.3 y A.4 se muestran las vistas generales y modelado tridimensional digital en software para su fabricación en impresora 3d (deposición de material) con polímero ácido poliláctico o poliácido láctico (PLA). El modelado principal e impresiones se muestran en las figuras A.5 y A.6. La estructura principal se muestra en la figura A.7 y está fabricada en el plástico acrílico utilizando corte láser a control numérico.

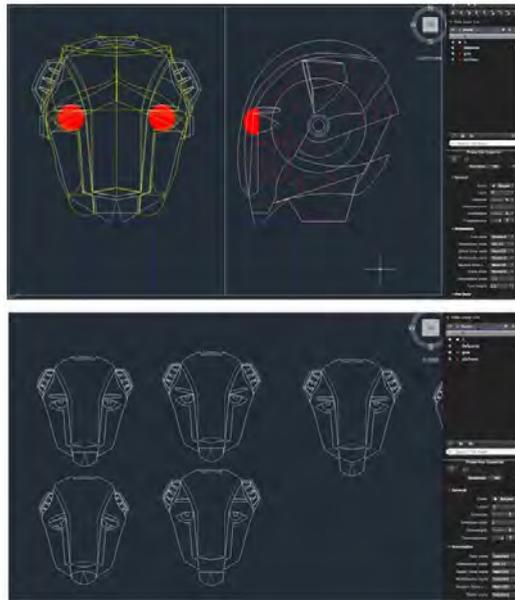


Figura A.3: Vistas generales del prototipo facial inicial.

En la siguiente dirección se muestra un vídeo con la prueba de funcionamiento <https://www.youtube.com/watch?v=Na2aH8mgEgk>. El modelo fue modificado a partir de las necesidades del proyecto y uso de equipo especializado, lo que alteró radicalmente el concepto inicial. El cambio se muestra en la figura A.8.

La figura A.9 presenta la estructura para el montaje del rostro en Golem-III; el diseño respondió a una adaptación más que a un concepto original los cuál desfavoreció a su impacto en el usuario y expectativas del grupo Golem. El diseño del rostro final fue gracias al trabajo colaborativo con el Grupo y al cambio de posición de los elementos funcionales.

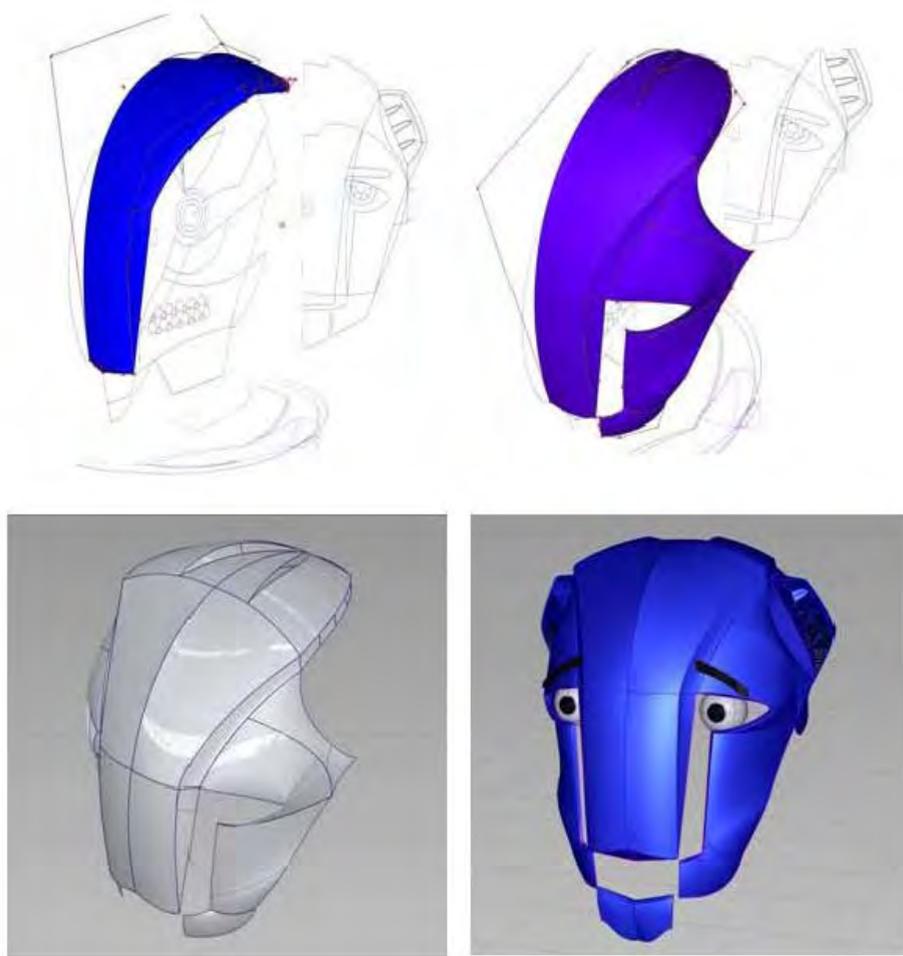


Figura A.4: Modelado tridimensional digital del prototipo facial inicial.



Figura A.5: Fabricación del modelo físico en impresión tridimensional.



Figura A.6: Cubierta fabricada con impresión tridimensional.



Figura A.7: Estructura para montaje de impresión tridimensional.

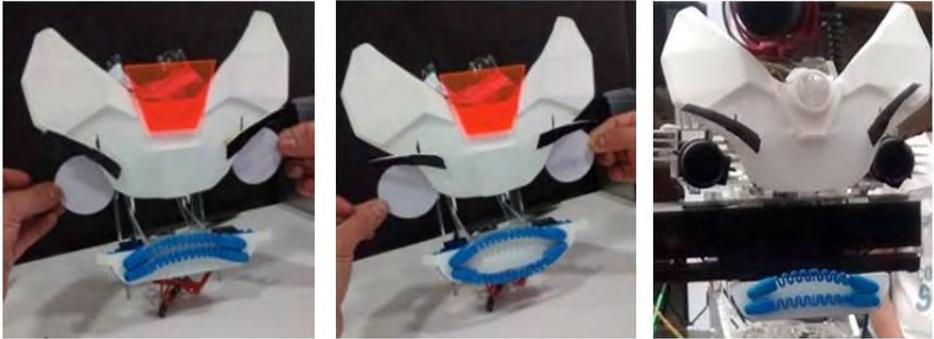


Figura A.8: Modificación del diseño inicial y propuesta para montaje en Golem-III

El desarrollo se realizó a partir de diseños preliminares y generación de piezas únicas como modelos de exploración. El diseño del rostro final se presenta en la figura A.11; se muestra la distribución de elementos funcionales con mejor organización y distribución, sin afectar la funcionalidad básica del robot.

El diseño del rostro final fue gracias al trabajo colaborativo con el Grupo y al cambio de posición de los elementos funcionales. El desarrollo se realizó a partir de diseños preliminares y generación de piezas únicas como modelos de exploración (figura A.10). El diseño del rostro final se presenta en la figura A.11; se muestra la distribución de elementos funcionales con mejor organización sin afectar la funcionalidad básica del robot.



Figura A.9: Estructura para el rostro de Golem-III



Figura A.10: Modelo de exploración realizado en madera

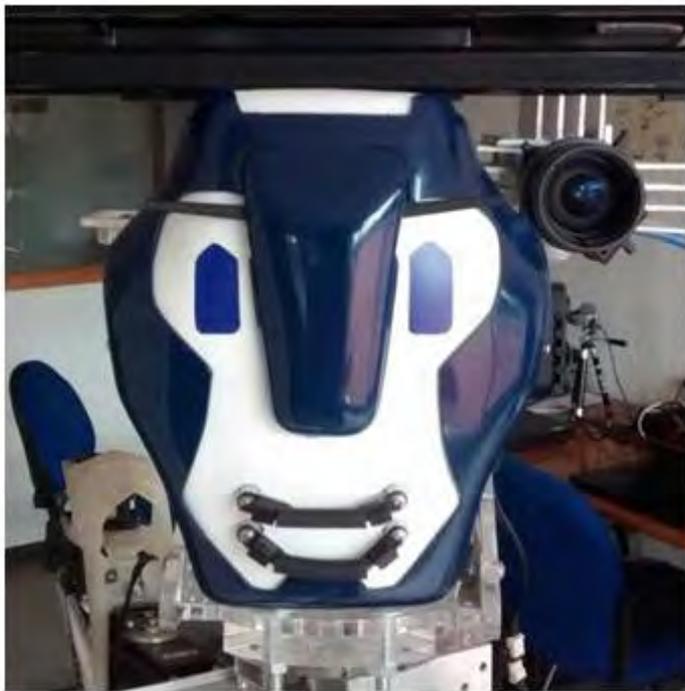


Figura A.11: Prototipo facial final de Golem-III