

Aplicación de modelos de emociones a agentes

Estudio de modelos y arquitecturas de emociones para su aplicación en sistemas de IA en general y multiagentes en particular.

Autores:

José Manuel Llorens Ripollés
Jaime Juan Pérez

Bajo la supervisión de:

Vicente Juan Botti Navarro

Trabajo presentado para la asignatura:

Agentes inteligentes

Departamento: **D. Sist. Informáticos Y Computación**

Universidad: **Universitat Politècnica de Valencia**

Fecha: **Mayo 2017**

Contents

1	Abstract	2
2	Modelo	2
2.1	Introducción	2
2.2	El modelo OCC	2
2.3	Ambigüedades y clarificaciones	4
2.4	Revisión del Modelo OCC	5
3	Arquitectura	6
3.1	Introducción	6
3.2	Arquitectura Wasabi	7
3.2.1	Emociones	7
3.2.2	Humor	8
3.3	Emociones en WASABI	8
3.3.1	Nueve emociones primarias	8
3.3.2	Emociones secundarias	8
3.3.3	Grado de conciencia de las emociones primarias	10
3.3.4	Grado de conciencia de las emociones secundarias	10
3.3.5	Conexión cuerpo mente	10
3.4	Aplicación en el juego Skip-Bo y primer estudio	11
3.4.1	Aplicación en el juego Skip-bo	11
3.4.2	Primer estudio	13
4	Conclusión	14

1 Abstract

Esta investigación está enfocada al estudio de modelos de emociones en agentes, práctica que se ha extendido por distintos campos como el de la robótica o el de la psicología.

Lo primero que definiremos es un modelo para formalizar las emociones y ser así capaces de aplicarlas a los agentes. Aunque existen varios modelos psicológicos en los que basarnos, nosotros nos basaremos en el modelo OCC definido por Ortony, Clore Collins[3]

La implementación de estos modelos en los agentes se realiza mediante arquitecturas, las cuales reciben estímulos que son tratados siguiendo el modelo, generando una emoción que el agente será encargado de representar. Cada arquitectura sigue unas reglas distintas pero generalmente todas tienden a hacer una diferencia entre en emociones primarias y secundarias. En nuestro caso nos centraremos en la arquitectura WASABI[2]

Finalmente presentaremos algunos ensayos realizados sobre la arquitectura WASABI y estudiaremos los resultados obtenidos en estos.

Una resumen multimedia de este trabajo puede ser visualizado en:
<https://www.youtube.com/watch?v=nRHHLpOMzmAfeature=youtu.be>

2 Modelo

2.1 Introducción

Las emociones son inherentemente subjetivas y difíciles de definir y estructurar. Los mismos estímulos pueden causar diferentes reacciones en diferentes humanos, de manera que, hasta cierto punto, las emociones se pueden clasificar como no deterministas. No obstante para poder hacer uso de ellas en el ámbito de la computación, necesitaremos tenerlas claramente definidas y estructuradas. Por lo tanto necesitamos definir un modelo subyacente que modele las emociones y las adecue para su uso por un agente. La idea principal de un modelo radica en que existen unos principios neuronales básicos que son la base de las emociones. La explosión de la inteligencia artificial ha hecho patente la necesidad de integrar emociones a los diferentes agentes y, gracias a esto, el estudio de modelos se ha visto incrementado en los últimos años. Uno de los modelos más extendidos es el Modelo OCC [3]

2.2 El modelo OCC

El modelo OCC, uno de los mas ampliamente usados, define y estructura 22 emociones en una jerarquía dividida en 3 ramas(Figura 1). Estas 3 ramas divergen en el causante de la emoción, sea este un evento, otro agente o el aspecto de un objeto. Debido a que en el campo de los agentes también es ampliamente usada la división eventos, agentes, objetos, el Modelo OCC se presenta adecuado para su uso en la aplicación de emociones. Una descripción de las 22 emociones puede ser encontrada en la tabla 1.

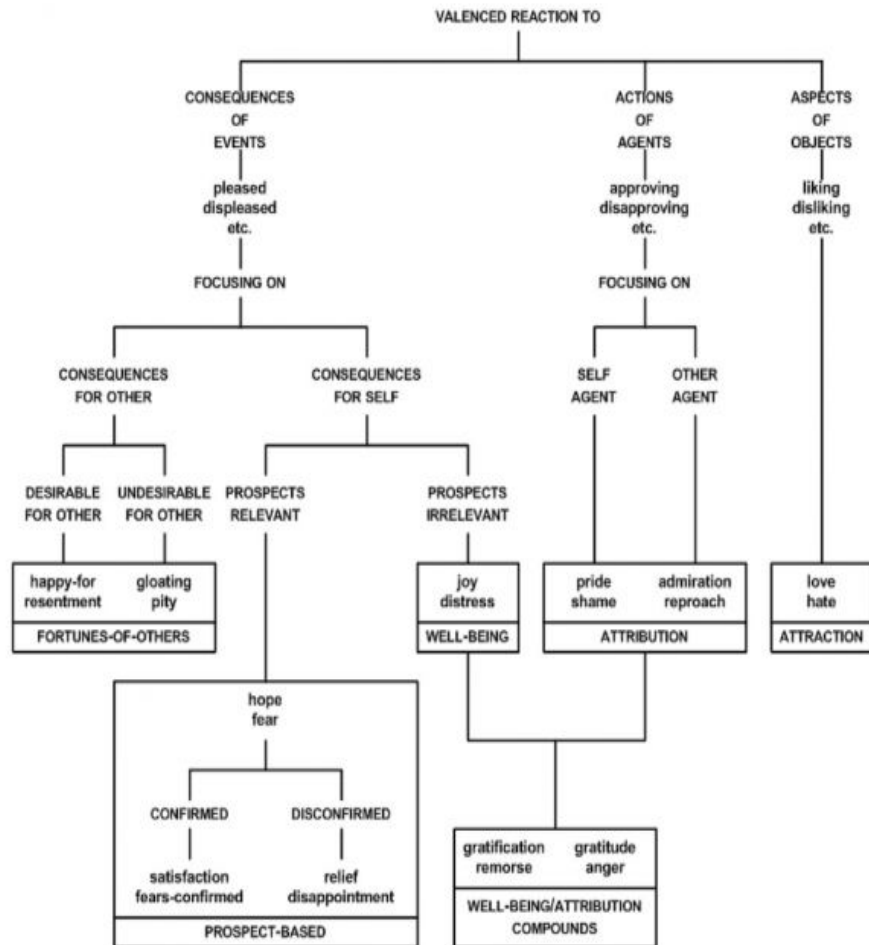


Figure 1: The OCC Mode

Alegría:	(Satisfecho sobre) un evento deseable
Angustia:	(Insatisfecho sobre) un evento indeseable
Feliz-por:	(Satisfecho sobre) un evento presumiblemente deseable por otra persona
Pena:	(Insatisfecho sobre) un evento presumiblemente indeseable por otra persona
Regodeo:	(Satisfecho sobre) un evento presumiblemente indeseado por otra persona
Resentimiento:	(Insatisfecho sobre) un evento presumiblemente deseable por otra persona
Esperanza:	(Satisfecho sobre) la perspectiva de un evento deseable
Miedo:	(Insatisfecho sobre) la perspectiva de un evento indeseable
Satisfacción:	(Satisfecho sobre) la confirmación de la perspectiva de un evento deseable
Confirmación de miedo:	(Insatisfecho sobre) la confirmación de la perspectiva de un evento indeseable
Alivio:	(Satisfecho sobre) la desconfirmación de la perspectiva de un evento indeseable
Decepción:	(Insatisfecho sobre) la desconfirmación de la perspectiva de un evento indeseable
Orgullo:	(Aprobatorio de) las acciones loables de uno mismo
Vergüenza:	(No aprobatorio de) las acciones culpables de uno mismo
Admiración:	(Aprobatorio de) las acciones loables de otra persona
Reproche:	(No aprobatorio de) las acciones culpables de otra persona
Gratificación:	(Aprobatorio de) las acciones loables de uno mismo y (Estar satisfecho sobre) el evento deseable consecuente.
Remordimiento:	(No aprobatorio de) las acciones culpables de uno mismo y (Estar insatisfecho sobre) el evento indeseable consecuente
Gratitud:	(Aprobamiento de) las acciones loables de otra persona y (Estar satisfecho de) el evento deseable consecuente
Ira:	(No aprobatorio de) las acciones culpables de otra persona y (Estar insatisfecho de) el evento indeseable consecuente
Amor:	(Gusto por) un objeto atractivo
Odio:	(Desagrado por) un objeto no atractivo

Tabla 1:Tipos de emociones en el modelo OCC

2.3 Ambigüedades y clarificaciones

La presentación de la estructura del modelo OCC se hace de forma en que se pueda asumir como un diagrama de herencia. Aun así, al intentar usar ese diagrama en el campo de la ciencia de computadores se revelan ciertas ambigüedades que pueden impedir su uso. Algunas de esas ambigüedades son:

1. En la tabla 1 se puede apreciar que se hace un uso extensivo de la frase "Evento deseable". No obstante no es el evento en si lo que es deseable, sino sus consecuencias. Por lo tanto evento deseable debe ser leído como "consecuencias deseables de un evento"

2. Cada una de las tres ramas principales esta encabezada por un set de palabras emocionales contrarias entre si, por ejemplo: aprobación/desaprobación. Estos tres pares de palabras se suponen los menos diferenciables emocionalmente. Por ejemplo: "Estar satisfecho" representa una situación donde el agente ha evaluado la consecuencia de un evento como deseable, no obstante no se dice nada de si esa consecuencia es vista como una prospección o como actual, por lo que no se puede diferenciar entre esperanza o alegría.
3. La alegría esta clasificada como una emoción lanzada por un evaluamiento deseable de las consecuencias de un evento, estando el foco en las consecuencias para uno mismo. No obstante en la especificación de alegría esta es dada como "Satisfecho sobre un evento deseable", sin ningún tipo de referencias sobre si estas se centran en uno mismo o no.
4. La alegría se define como estar satisfecho sobre un evento deseable, mientras que la angustia se define como estar insatisfecho sobre un evento indeseable. Cruzando estas dos especificaciones uno se puede preguntar si se puede estar satisfecho sobre un evento indeseable, o insatisfecho sobre un evento deseable. Esta cruce puede introducir ambigüedades, por ejemplo con "esperanza", los términos "Satisfecho sobre" se centran en el evento deseable o en tener el prospecto del mismo?
5. Considerando el grupo de emociones de "Feliz por otros" la consecuencia que es deseable para la otra persona, también debe ser considerada deseable para uno mismo hasta cierto punto. Dado que la consecuencia también es deseable para uno mismo, las condiciones de "alegría" se cumplen. De tal forma, tanto alegría como angustia son generalizaciones de resentimiento, feliz-por, etc, y como tal deberían estar representados como padres de ellas en la jerarquía.

2.4 Revisión del Modelo OCC

El modelo OCC se presenta como un diagrama de herencia, en el que cada emoción es como su padre más una esencialización. No obstante, el uso de el modelo OCC en el campo informático presenta una serie de ambigüedades en la estructura lógica subyacente (Como hemos descrito en el apartado anterior). Por lo tanto nosotros nos centraremos en el estudio de la revisión propuesta en "The OCC Model Revisited"[1]. La estructura resultante se puede observar en la figura [2]. Como podemos ver se han introducido ciertos cambios como:

1. Se define la jerarquía de una manera explicita.
2. Sobre cada conexión están escritas las condiciones adicionales que hacen de la emoción una especialización de sus padres.
3. Satisfacción, miedos-confirmados, alivio y decepción se han movido debajo de esperanza/miedo, convirtiéndose así en especializaciones de alegría/angustia.

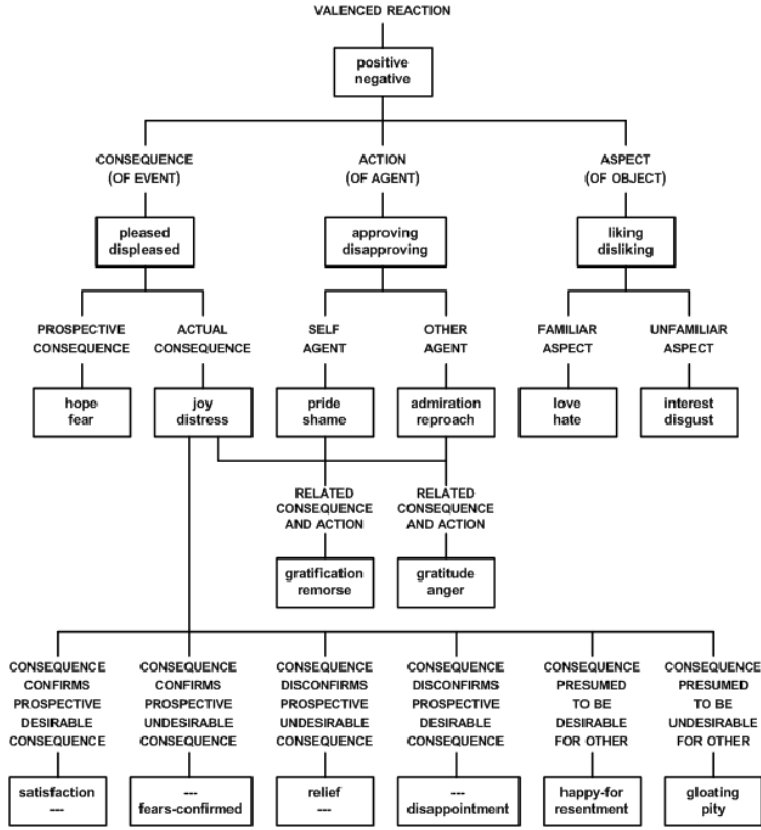


Figure 2: The OCC Model Revisited

4. Feliz-por, resentimiento, regodeo y pena han sido puestos en la parte mas baja-derecha de la figura, esto se debe al punto 5 del apartado anterior.
5. Las descripciones ahora pueden ser fácilmente formadas siguiendo la conexión del hijo al nodo padre.

3 Arquitectura

3.1 Introducción

Una arquitectura es un sistema artificial que reproduce ciertos aspectos de un sistema natural, que puede ayudarnos a entender los mecanismos internos que conducen a efectos particulares. Centrándonos en los sistemas naturales humanos, los estudios realizados sobre la materia han concluido que podemos diferenciar dos comportamientos, emociones primarias y secundarias.

Las emociones primarias son innatas, es decir, son naturales del sistema y son desarrolladas durante el desarrollo evolutivo para apoyar reacciones rápidas en caso de peligro inmediato; son el tipo de emociones que podríamos encontrar en un niño de un año de edad. Finalmente, son llamadas como el “Core Affect” ya que forman la base del resto de emociones. Por otro lado, las emociones secundarias son fruto de un proceso cognitivo más elevado, basado en la habilidad de evaluar preferencias entre resultados y expectativas. Este tipo de emociones es adquirido durante el proceso de aprendizaje en el contexto social e influyen las expresiones corporales a través de los mismos mecanismos que las emociones primarias, es por esto que están formadas por emociones primarias.

3.2 Arquitectura Wasabi

La arquitectura WASABI combina dinámicas de emociones corporales con valoraciones cognitivas para simular emociones primarias infantiles así como emociones secundarias adultas, más elaboradas cognitivamente. Se basa en la simulación de una dinámica de emoción en un espacio emocional tridimensional y en la aplicación sobre el humano virtual MAX.

El espacio tridimensional es el conocido como PAD, el cual usa las dimensiones placer, excitación y dominio (Pleasure, Arousal and Dominance) para representar todas las emociones. WASABI hace que una conceptualización sobre emociones y humor, que veremos a continuación.

3.2.1 Emociones

Las emociones son entendidas como estados con una cualidad e intensidad específicas, resultado de un complejo proceso neuropsicológico para comunicarse. Estos procesos incluyen actividad neuronal y respuestas del cuerpo. Las emociones se clasifican en primarias y secundarias, pero cada emoción tiene su balance positivo o negativo en función un valor, y en comparación con el humor, las emociones duran menos. Las principales diferencias entre emociones primarias y secundarias son las enumeradas a continuación:

- Las emociones secundarias están basadas en estructuras de datos más complejas que las primarias.
- Las emociones secundarias son más dependientes de las habilidades cognitivas de razonamiento del agente.
- Los disparadores de las emociones secundarias son aprendidos basándose en la conexión de las emociones primarias con eventos, agentes y objetos.
- El agente expresa que es consciente de sus emociones secundarias verbalmente.
- Las expresiones faciales de emociones primarias pueden acompañar a las emociones secundarias.

	Primary emotion	Facial expr. (Ekman)	PAD values	Base intensity i_{pe}
1	Angry	Anger (<i>anger</i>)	(80, 80, 100)	0.75
2	Annoyed	Sad (<i>sadness</i>)	(-50, 0, 100)	0.75
3	Bored	Bored (<i>none</i>)	(0, -80, 100)	0.75
4	Concentrated	Neutral (<i>none</i>)	(0, 0, ± 100)	0.75
5	Depressed	Sad (<i>sadness</i>)	(0, -80, -100)	0.75
6	Fearful	Fear (<i>fear</i>)	(-80, 80, 100)	0.25
7	Happy	Happy (<i>happiness</i>)	(80, 80, ± 100)	0.75
			(50, 0, ± 100)	0.75
8	Sad	Sad (<i>sadness</i>)	(-50, 0, -100)	0.75
9	Surprised	Surprised (<i>surprise</i>)	(10, 80, ± 100)	0.0

Figure 3: Tabla de las diferentes emociones primarias sobre el espacio PAD

3.2.2 Humor

El humor es modelado como estado de fondo con una calidad afectiva mucho más simple que las emociones. Es modelado como un sentimiento de bienestar general del agente en una escala bipolar de balance positivo contra negativo. La interconexión entre humor y emociones resulta en una dinámica de emociones, donde el humor influencia la obtención de emociones de una forma en que se consigue congruencia entre humor y emociones.

3.3 Emociones en WASABI

WASABI representa nueve emociones primarias y tres emociones secundarias, empleando PAD para calcular estas.

3.3.1 Nueve emociones primarias

Las emociones primarias son estados afectivos innatos, disparados por reflejos en caso de amenaza. Cada una de las emociones primarias está situada en el espacio de PAD según la figura 3

Basándose en la distancia sobre el espacio de emociones, las emociones son dinámicamente obtenidas según la cognición del agente, la cual se encarga también de disparar estas emociones de forma que su intensidad base sea 1. En la Figura 4 podemos ver siete expresiones faciales de MAX (humano virtual mencionado anteriormente).

3.3.2 Emociones secundarias

Recordando el modelo OCC, las emociones basadas en promesas (un grupo del modelo de emociones OCC) se consideran que pertenecen a las emociones secundarias, porque su proceso de valoración incluye la evaluación de eventos contra experiencias y expectativas. De este grupo WASABI es capaz de simular esperanza, miedos confirmados y alivio, y a continuación explicaremos su estructura.

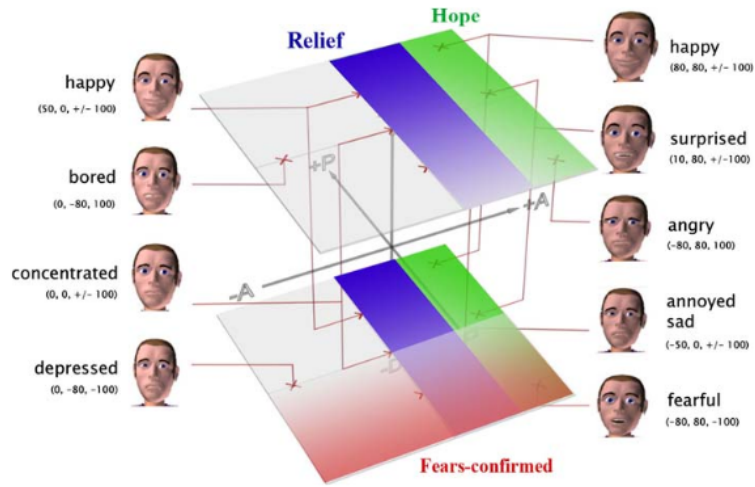


Figure 4: Expresiones faciales del humano virtual MAX

Esperanza La esperanza es descrita como el resultado de la valoración de un evento futuro, si el evento es deseable para uno mismo se dice que es probable ser satisfecho por la posibilidad de un evento deseable. La esperanza se representa en PAD como:

- Placer: cuanto más placer siente es más probable que un agente sea consciente de la esperanza.
- Excitación: la esperanza es más probable de obtenerla cuanto más alta sea la excitación del agente.
- Dominio: la probabilidad del agente de ser consciente de la esperanza se modela independientemente del nivel general de dominio.

Miedos confirmados Los miedos confirmados aparecen cuando se está descontento sobre la confirmación de un futuro evento no deseable. Estos son representados en PAD como:

- Placer: la probabilidad de ser consciente de los miedos confirmados aumenta cuanto más descontento se sienta el agente.
- Excitación: se asume que son independientes de la excitación.
- Dominio: los miedos confirmados solo son percibidos por el agente cuando este se siente sumiso.

Alivio El alivio se describe como tener experiencia en situaciones que son satisfactorias por la desconfirmación de un evento futuro no deseable. Se representa en PAD como:

- Placer: es más probable que un agente sienta alivio cuanto más placer sienta.
- Excitación: solo en caso de niveles bajos de excitación el agente es capaz de sentir alivio.
- Dominio: la probabilidad de ser consciente de alivio se considera independiente del estado de dominio del agente.

3.3.3 Grado de conciencia de las emociones primarias

El grado de conciencia en las emociones primarias viene dado por la distancia de la emoción misma respecto del valor PAD. Cuanto menor sea esta distancia en mayor grado estará presente la emoción. Se establece cierto umbral de distancia ϕ_{pe} a partir del cual la intensidad de la emoción será calculada de la siguiente manera:

$$\omega_{pe} = \left(1 - \frac{d_{pe} - \Delta_{pe}}{\Phi_{pe} - \Delta_{pe}}\right) \mid \phi_{pe} > \delta_{pe} \forall pe \in pe_1, \dots, pe_9$$

- ω_{pe} : el grado de intensidad resultado.
- d_{pe} : distancia entre el valor PAD actual y el valor de la emoción primaria.
- ϕ_{pe} : cierto umbral de intensidad.
- δ_{pe} : umbral distancia por debajo del cual la intensidad de la emoción se equipara a la intensidad base.

3.3.4 Grado de conciencia de las emociones secundarias

Los valores actuales del PAD también son usados para calcular el grado de las tres emociones secundarias. Las emociones secundarias, en lugar de las primarias tendrán las intensidades base como 0. Por este motivo las emociones secundarias necesitarán ser "lanzadas" por un proceso cognitivo para hacerse conscientes en el agente. Dado que las emociones secundarias tienen una intensidad base de 0 necesitaremos definir una función que defina como decrecerá la intensidad una vez lanzadas, asegurándonos así que vuelven a su nivel base. Esto se consigue mediante el parámetro "lifetime" que será propio de cada emoción (10.0 como base) y la función de decrecimiento (linear como base)

3.3.5 Conexión cuerpo mente

Para introducir este punto nos basaremos en la figura[5] en la que observamos los procesos que se llevan a cabo durante el procesamiento de emociones y que definen la encarnación del agente (en nuestro caso el agente MAX, que podemos visualizar en el fondo de la imagen).

Nuestro agente recibe una serie de estímulos que pueden ser internos o externos; estos impulsos pueden ser valorados inconscientemente o conscientemente.

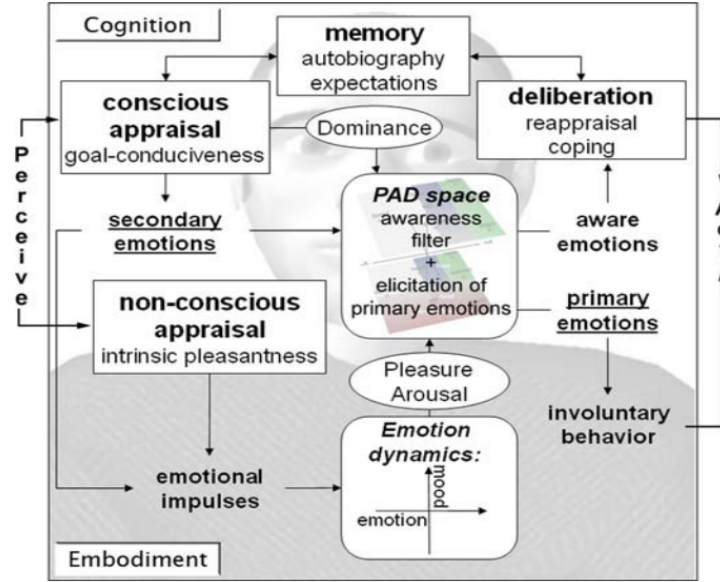


Figure 5: Distinción cognición-personificación en la arquitectura WASABI

La valoración inconsciente se realiza mediante el envío de un impulso emocional al módulo encargado de las dinámicas de la emoción, estableciendo una vía rápida para el lanzamiento de emociones primarias. Por otro lado, tenemos la valoración consciente, en la cual MAX usará sus habilidades de razonamiento para actualizar su memoria y generar expectativas, además de derivar su nivel de dominio. En esta etapa, se generan emociones secundarias que también son mandadas como impulsos emocionales.

Las emociones secundarias son filtradas en el espacio PAD antes de que MAX sea consciente de ellas. Independientemente del filtrado, la dinámica de emociones es modificada en función de los impulsos emocionales que recibe, modulando los valores de placer y excitación.

Finalmente, después del filtrado se obtienen las emociones primarias, que se reflejarán con comportamientos involuntarios (cambios en las expresiones faciales), y una consciencia de que existen emociones (secundarias) que mediante un proceso de deliberación llevarán a realizar ciertos actos.

3.4 Aplicación en el juego Skip-Bo y primer estudio

3.4.1 Aplicación en el juego Skip-bo

En este apartado se extenderán las capacidades cognitivas de MAX mediante una implementación del juego Skip-Bo.

Observando la figura vemos que el escenario está formado por tres agentes, el agente BDI, el Emotion-Agent y el agente de visualización. Estos módulos,

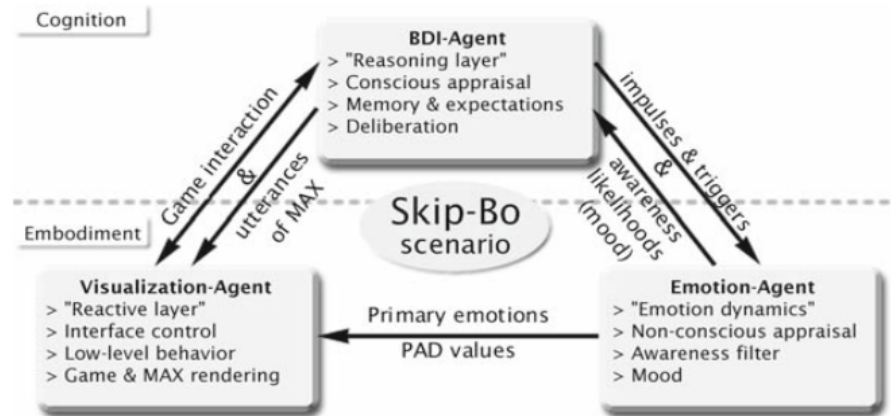


Figure 6: Los tres agentes software mas importantes y sus interconexiones respecto al paso de mensajes.

a grandes rasgos, se relacionan de forma que el Emotion-Agent recibe impulsos emocionales del agente BDI y es actualizado según sea consciente de las emociones. Por otro lado, el agente BDI también calcula cuando MAX es dominante y junto a las emociones según el espacio PAD alimenta al agente de visualización.

A continuación veremos un ejemplo de funcionamiento de lo anteriormente dicho, basándonos en el diagrama de la figura[7]:

Los tres agentes están representados por los tres rectángulos que podemos ver en la parte superior del diagrama. Dentro de cada rectángulo tenemos los distintos planes que intervendrán, excepto en el caso del agente de visualización que solo se encarga de visualizar, aunque en la realidad también recoja información del jugador humano.

Siguiendo una secuencia temporal de abajo arriba, el primer plan que se activa es el de generar una expectativa, este manda un impulso negativo al Emotion-Agent el cual actualizara el espacio PAD. Paralelamente a esto, el mismo plan dispara la emoción primaria de miedo, dado que MAX espera que la siguiente carta jugada por el adversario sea importante; sin embargo, aunque esto no es suficiente para disparar la emoción de miedo, se manda la señal de “un poco miedoso” al agente de visualización y esto se debe a la naturaleza de la emoción miedo la cual va incrementado según una escala.

El siguiente plan que se activa es el de comprobar una expectativa, para ver si la acción que ha tomado el humano se corresponde con lo esperado. En este caso sí que se corresponde y el plan activa una señal de miedos confirmados que llega al Emotion-Agent, donde se maximiza su valor base y la probabilidad de darse cuenta de los miedos confirmados deja de ser nula. Todo esto se transmite al agente de visualización y se encarga de expresarlo verbalmente.

Un humano daría la siguiente descripción del caso que acabamos de comentar:

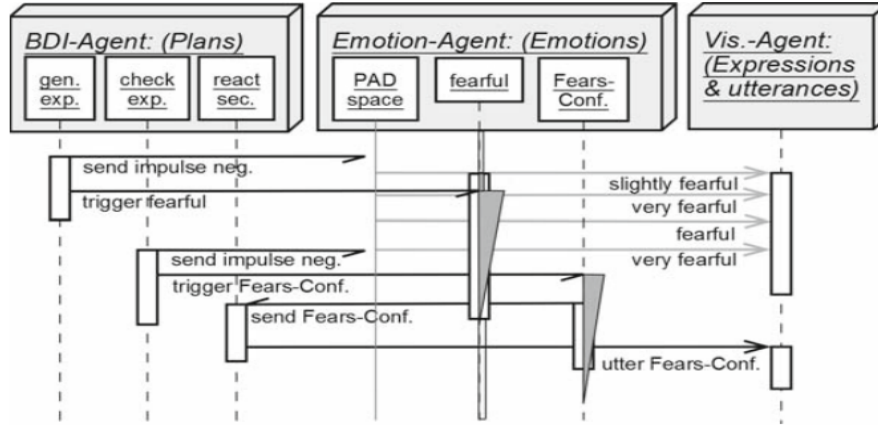


Figure 7: Diagrama de frecuencia del flujo de información entre los agentes software

Cuando ha acabado su turno, MAX se ha dado cuenta de que podía jugar una carta importante, esto lo he notado porque su expresión facial ha cambiado y parecía algo asustado, después de jugar la carta me ha admitido que se había asustado.

3.4.2 Primer estudio

El experimento parte de la hipótesis de que si MAX emplea emociones primarias y secundarias, al juzgar la edad se le considerará más mayor que si usará emociones primarias.

Referente a la hipótesis, la prueba tiene un diseño que se divide en dos casos:

1. Aquel en el MAX solo simula emociones primarias y las expresa con expresiones faciales y sonidos. MAX reconoce las jugadas del oponente verbalmente y las valora negativamente, mientras que su progreso en el juego lo valora positivamente. Finalmente, cuando el oponente cometa un fallo MAX entrará en estado dominante y le corregirá, de lo contrario permanecerá sumiso.
2. Inclusión tanto de emociones primarias como secundarias, estas últimas son simuladas junto a la configuración del caso uno y son expresadas verbalmente por MAX en caso de que la probabilidad de consciencia sea positiva.

En cuanto a la forma de llevar a cabo el experimento, primero se realiza una familiarización de los participantes con el entorno. Para ello se dispone de un agente MAX sin emociones, con el que los participantes aprenden las reglas del juego. Una vez terminado el proceso de introducción el participante tendrá que

jugar una partida con el agente MAX, esta vez con emociones, y al terminar deberá realizar un cuestionario.

Una vez se realizó el experimento los resultados fueron que el agente parecía más joven cuando solo empleaba emociones primarias que cuando las usaba junto a emociones secundarias.

4 Conclusión

En esta investigación sobre las emociones en agentes inteligentes nos hemos centrado principalmente en el Modelo OCC y en la arquitectura WASABI. Según lo desarrollado el modelo OCC es un buen punto de partida para realizar investigaciones y pruebas acerca del campo. Pese a ser un modelo que presenta ciertas ambigüedades, si es tratado de la forma adecuada y conociendo sus límites se puede sacar un gran rendimiento de él. Prueba de esto anterior es la arquitectura WASABI, que basándose en el modelo OCC desarrolla todo un sistema emocional con el cual consigue dotar a un agente de características emocionales humanas. Para desarrollar este proceso, la arquitectura hace una división entre emociones primarias y secundarias, implementando un sistema que unificará los distintos módulos y realizará los cálculos correspondientes para generar la emoción correspondiente ante la llegada de un estímulo.

References

- [1] Mehdi Dastani Bas R. Steunebrink and John-Jules Ch. Meyer. *The OCC Model Revisited. (English)*.
- [2] Ipke Wachsmuth Christian Becker-Asano. *Affective computing with primary and secondary emotions in a virtual human. (English)*.
- [3] Andrew Ortony Gerald L. Clore Allan Collins. *The Cognitive Structure of Emotions. (English)*.