Modelado Causal en Marketing mediante Aprendizaje no Supervisado de Reglas de Asociación Difusas

Albert Orriols-Puig[1]

Jorge Casillas[2]

Francisco J. Martínez-López[3]

[1]Grup de Recerca en Sistemes Intel·ligents, Enginyeria i Arquitectura La Salle, Universitat Ramon Llull

[2]Dpto. Ciencias de la computación e Inteligencia Artificial, Universidad de Granada

[3]Dpto. Comercialización e Investigación de Mercados, Universidad de Granada

Índice

1. Descripción del Problema

2. Reglas de Asociación

3. Descripción de Fuzzy-CSar

4. Metodología Experimental

5. Resultados

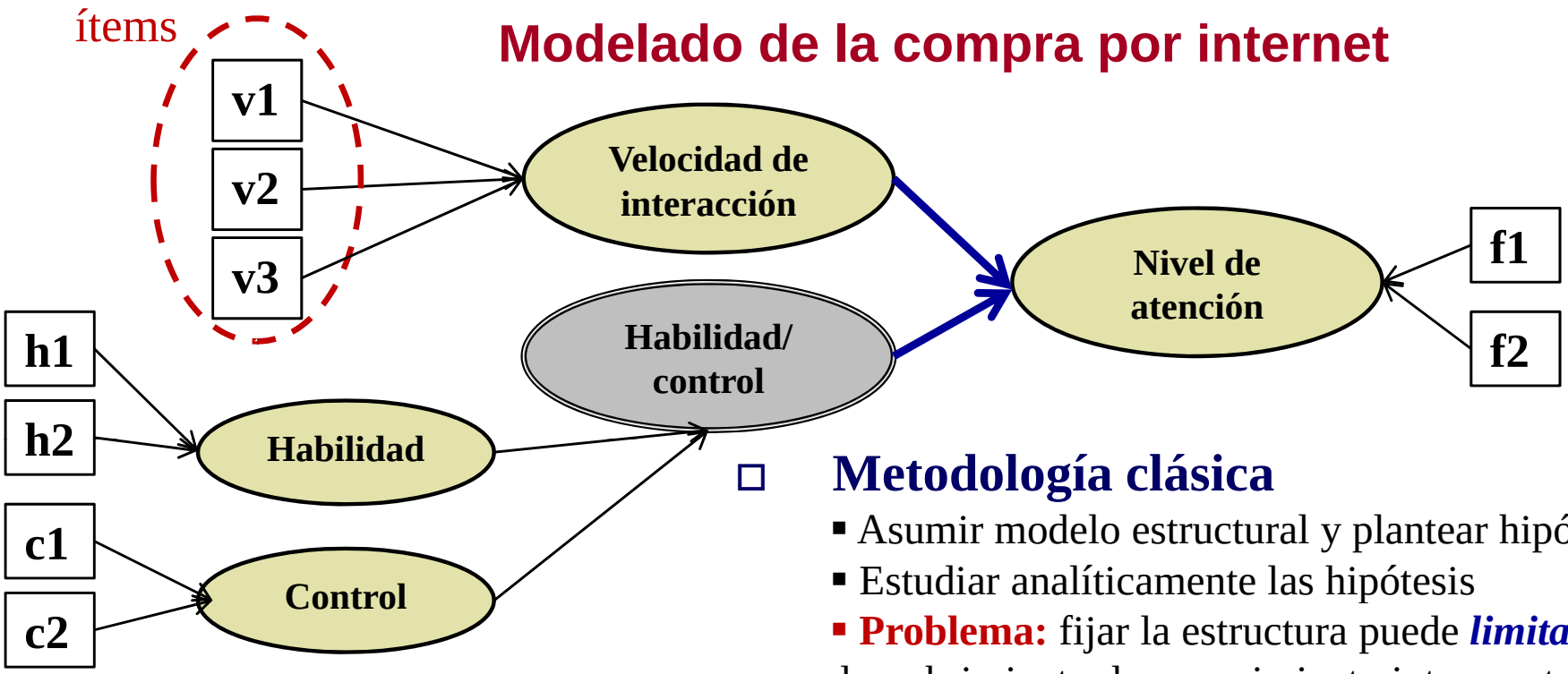
6. Conclusiones y Trabajo Futuro

Descripción del Problema

- **Búsqueda de oportunidades de mercado**
  - Sistemas de apoyo a la gestión de marketing
  - Extraer información clave para la toma de decisiones

- **Modelado del comportamiento del consumidor**
  - Definir las características o variables que se quieren medir
    - **Variables de primer orden**
    - **Variables de segundo orden definidas por varias variables de primer orden**
  - Datos recogidos de cuestionarios
    - **Preguntas indirectas para evaluar las variables**

Descripción del Problema



- **Metodología clásica**
  - Asumir modelo estructural y plantear hipótesis
  - Estudiar analíticamente las hipótesis
  - **Problema:** fijar la estructura puede ***limitar*** el descubrimiento de conocimiento interesante.

Velocidad de interacción	
v_1:	La interacción con las páginas Web es rápida y estimulante: {1,...,9}
v_2:	Internet es rápido: {1,...,9}
v_3:	Las páginas Web que visito usualmente se cargan lo suficientemente rápido: {1,...,9}

Cómo Afrontar el Problema?

- **Trabajo previo (Martínez & Casillas, 2008)**
  - Método multi-objetivo GCCL
    - **Fijar variable de salida**
    - **Extracción de reglas difusas**
  - Fuzzificación *multi-ítem* para tratar los datos
  - Permitió descubrir conocimiento clave no detectado por el método clásico

- **Propósito del trabajo actual**
  - Extraer reglas de ***asociación difusas*** con el objetivo de extraer relaciones entre variables que sean útiles y desconocidas para el experto en marketing
  - Comparar el nuevo sistema con el GCCL

Índice

1. Descripción del Problema

2. Reglas de Asociación

3. Descripción de Fuzzy-CSar

4. Metodología Experimental

5. Resultados

6. Conclusiones y Trabajo Futuro

Reglas de Asociación

- **Objetivo:** Extraer reglas que denoten relaciones entre variables con alto *soporte* y *confianza*

SI $X = \{x_1, x_2, \ldots, x_p\}$ **es** $\{F_{x1}, F_{x2}, \ldots, F_{xp}\}$ **ENTONCES**
$Y = \{y_1, y_2, \ldots, y_p\}$ **es** $\{G_{y1}, G_{y2}, \ldots, G_{yp}\}$

SI velocidad **es** *alta* **y** habilidad **es** *media* **ENTONCES**
nivel de atención **es** *alto*

- **Evaluación por soporte y confianza:**

$$\mathrm{sop(R)} = \frac{1}{N} \sum_{e=1}^{N} \mu_A(x^{(e)}) \cdot \mu_B(y^{(e)})$$

$$\mathrm{conf(R)} = \frac{\sum_{e=1}^{N} \left(\mu_A(x^{(e)}) \cdot \max\left\{1-\mu_A(x^{(e)}), \mu_B(y^{(e)})\right\}\right)}{\sum_{e=1}^{N} \mu_A(x^{(e)})}$$

Índice

1. Descripción del Problema

2. Reglas de Asociación

3. Descripción de Fuzzy-CSar

4. Metodología Experimental

5. Resultados

6. Conclusiones y Trabajo Futuro

Descripción de Fuzzy-CSar

- **Sistema clasificador de estilo Michigan**
  - Sistema de aprendizaje *online*
  - Uso de *algoritmos genéticos* para evolucionar el modelo
- **Representación del conocimiento**
  - Evolución de una *población* de *clasificadores*
  - Cada clasificador contiene
    - *Regla de asociación difusa* con *1 sola variable* en

$$\textbf{si } x_i \text{ es } \widetilde{A_i^k} \text{ y } \cdots \text{ y } x_j \text{ es } \widetilde{A_j^k} \textbf{ entonces } y_c \text{ es } \widetilde{B_c^k}$$

    - Conjunto de *parámetros*
      - *Soporte*
      - *Confianza*
      - *Fitness*
      - …

Donde *cada variable* está definida por varios ítems.

Fuzzificación multi-ítem.

Fuzzificación Multi-ítem

- **Objetivo: Unir la información de los ítems para determinar el grado de emparejamiento por cada variable**

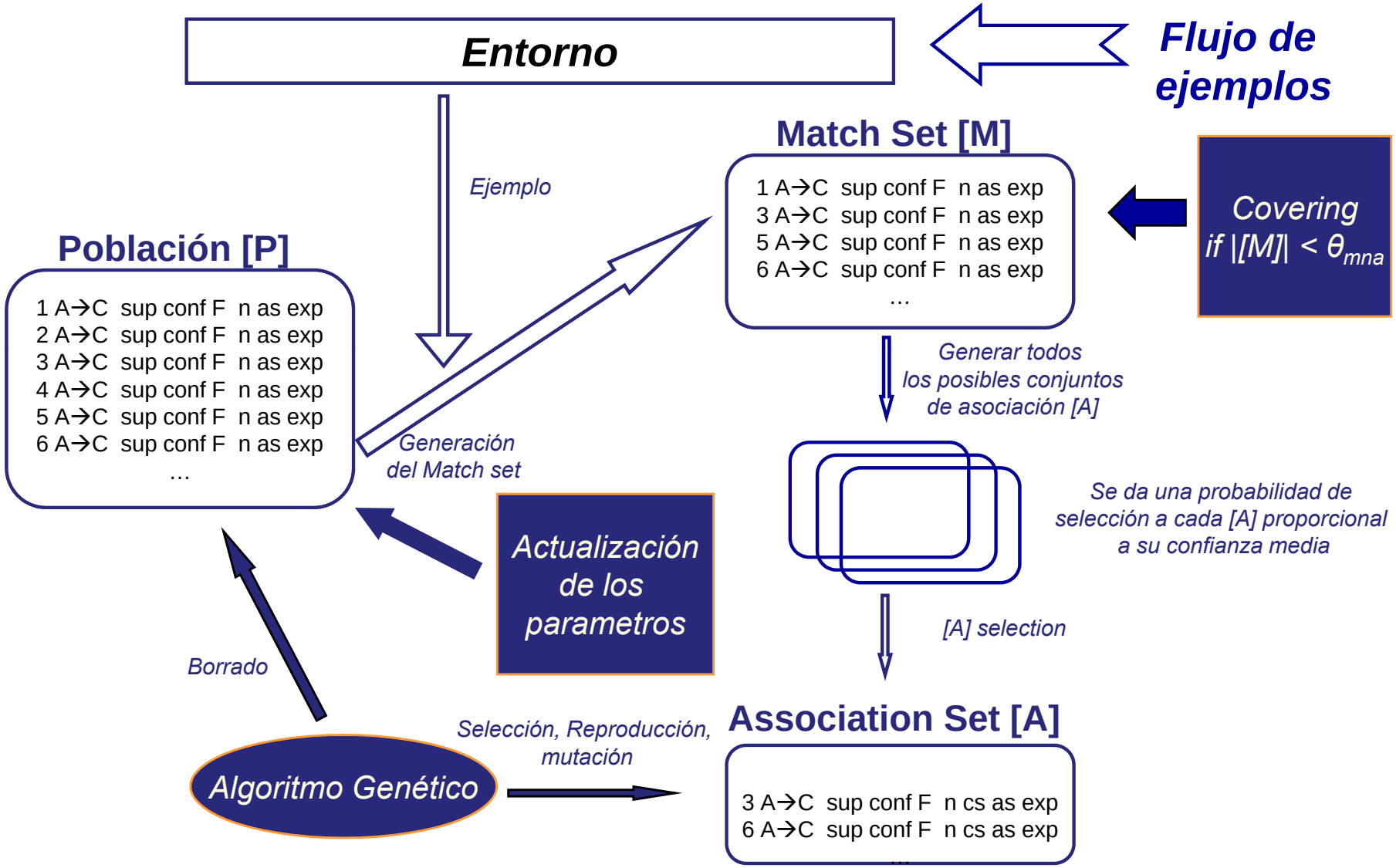
  - Computar la *T-conorma* grado de emparejamiento de la *variable i* con *cada ítem* de los p_i que representan la variable

$$\mu_{\widetilde{A}_i}(e_i) = max_{h_i=1}^{p_i} \mu_{\widetilde{A}_i}(x_i^{(e_i)})$$

  - *Variables de segundo orden* se computan como la *T-norma*

$$\mu_{\widetilde{A}_i}(e_i) = min_{k_i=1}^{s_i} \mu_{\widetilde{A}_i}(x_i^{(e_i)})$$

Proceso de Aprendizaje



Índice

1. Descripción del Problema

2. Reglas de Asociación

3. Descripción de Fuzzy-CSar

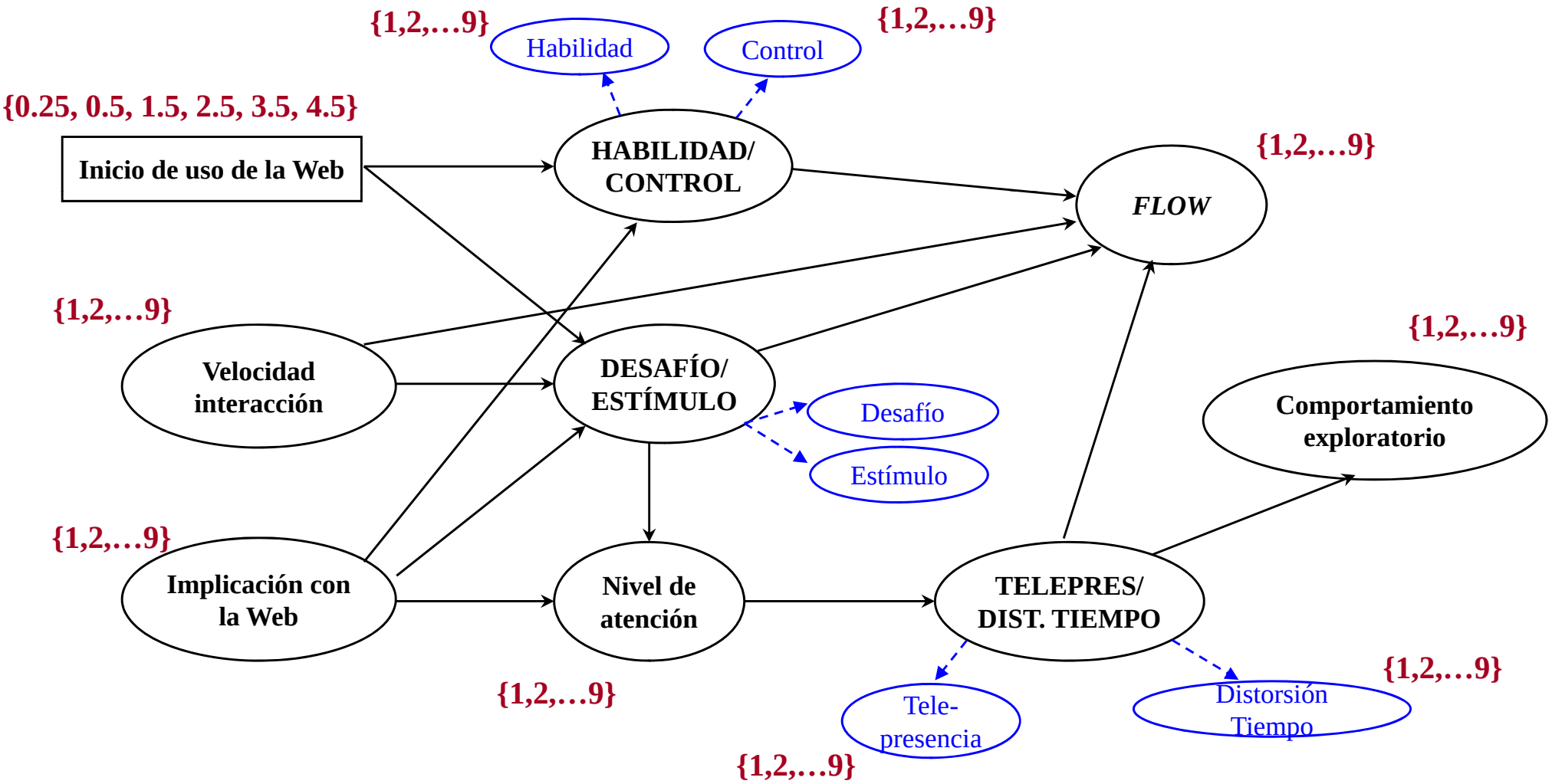
4. Metodología Experimental

5. Resultados

6. Conclusiones y Trabajo Futuro

Metodología Experimental

- Experimentación en el modelo causal presentado por Novak et al. (2000)



Metodología Experimental

- **Comparativa entre Fuzzy-CSar y GCCL basado en NSGA-II diseñado por Martínez & Casillas (2008)**

- **Experimento 1 (*4in1out*):**
  - 1 variable de salida: flow
  - 4 variables de entrada: habilidad/control, velocidad de interacción, desafío/estímulo y telepresencia/distorsión del tiempo

- **Experimento 2 (*8in1out*):**
  - 1 variable de salida: flow
  - Resto de variables como entrada
  - ¿Los resultados son similares a los obtenidos con 4 variables?

- **Experimento 3 (*9in9out*):**
  - Se permite cualquier variable en entrada o salida
  - ¿Se obtiene conocimiento interesante no detectado previamente?

Índice

1. Descripción del Problema

2. Reglas de Asociación

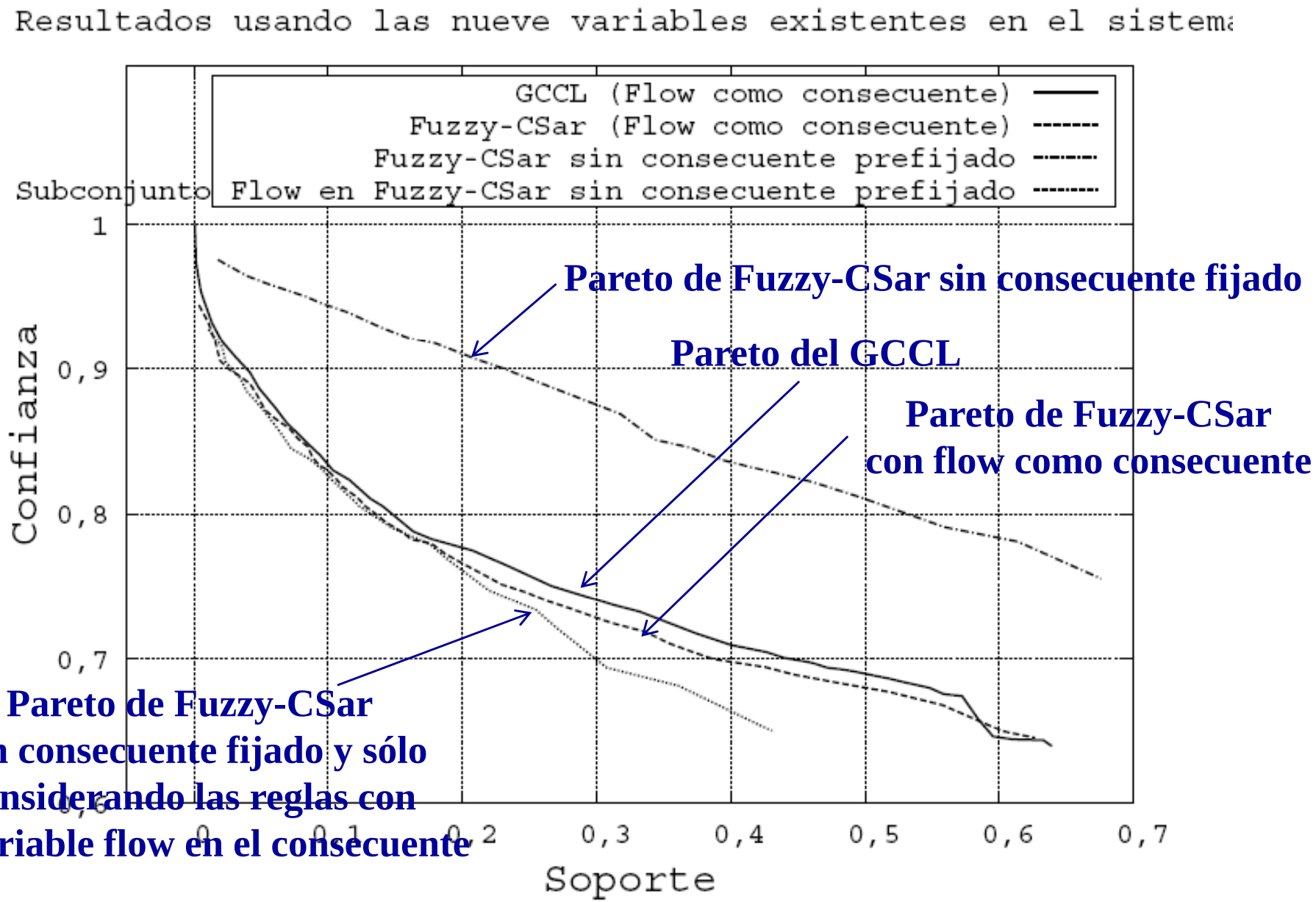
3. Descripción de Fuzzy-CSar

4. Metodología Experimental

5. Resultados

6. Conclusiones y Trabajo Futuro

Resultados: Análisis de los Paretos



Análisis del Valor de las Reglas

- **Obtención de reglas no contempladas en el modelo propuesto por Novak et al. (2000)**

SI implicación web es M y habilidad/control es P o M y nivelAtención es P o M y flow es P o M **ENTONCES** comportamientoExploratorio es M
Soporte = 0,22 y **conf** = 0,87

SI implicación web es P o M y desafío/estímulo es P o M y telepresencia/distancia de tiempo es M o L **ENTONCES** nivelAtención es M
Soporte = 0,21 y **conf** = 0,84

- **Utilidad de los resultados:**
  - Denotan relaciones no incluidas en el modelo estructural de Novak et al. (2000), proporcionando información útil
  - Permiten la revisión del modelo inicial

Índice

1. Descripción del Problema

2. Reglas de Asociación

3. Descripción de Fuzzy-CSar

4. Metodología Experimental

5. Resultados

6. Conclusiones y Trabajo Futuro

Conclusiones y Trabajo Futuro

- **Fuzzy-CSar es un sistema competente que:**
  - Aprende reglas de asociación sin tener que fijar los consecuentes de las reglas
  - Extrae reglas de asociación útiles y originales para los expertos de marketing
  - Obtiene frentes de Pareto similares a los extraídos por un sistema optimizador de frentes Pareto
  - Aprende online
    - **Grandes volúmenes de datos**
    - **Actualización del conocimiento frente cuestionarios nuevos**
- **Trabajo futuro:**
  - Estudio detallado de las capacidades de Fuzzy-CSar
  - Análisis de la diversidad de las reglas. Las reglas dominadas también pueden tener información útil

Modelado Causal en Marketing mediante Aprendizaje no Supervisador de Reglas de Asociación Difusas

Albert Orriols-Puig[1]

Jorge Casillas[2]

Francisco J. Martínez-López[3]

[1]Grup de Recerca en Sistemes Intel·ligents, Enginyeria i Arquitectura La Salle, Universitat Ramon Llull

[2]Dpto. Ciencias de la computación e Inteligencia Artificial, Universidad de Granada

[3]Dpto. Comercialización e Investigación de Mercados, Universidad de Granada