

Estudio del problema del matrimonio estable en un entorno multiagente

Estudiante: Alberto Hueso Alonso

Grado Ingeniería Informática
Especialidad Computación
Inteligencia Artificial

Consultor: David Isern Alarcón

Profesor responsable de la asignatura: Carles Ventura Royo

Fecha:07/01/2018

Introducción

- Intereses Personales

- Algoritmos de resolución de problemas específicos
- Sistemas multiagente
- Técnicas de aprendizaje computacional

TFG intenta conjugar los tres aspectos

- Elementos escogidos

- Algoritmo del matrimonio estable de Gale-Shapley
- JADE: Framework de software en Java para desarrollar SMA
- Weka: Software de aprendizaje computacional y minería de datos

- Objetivo: Averiguar si un clasificador de Weka puede anticipar los resultados del algoritmo del matrimonio estable de Gale-Shapley.

Problema del matrimonio estable

- Existen dos grupos de elementos que hay que emparejar
- Cada elemento ordena los elementos del otro grupo según sus preferencias para ser emparejado
- Las parejas pueden ser estables o inestables
- Las parejas son inestables si alguno de sus miembros prefiere a un miembro de otra pareja que, a su vez, le prefiere a él antes que a su propia pareja. En caso contrario, son estables

Matrimonio estable

Ejemplo

Kepa → Judith, Montse, Eva	Eva → Antonio, Kepa, Jordi
Jordi → Montse, Judith, Eva	Montse → Jordi, Antonio, Kepa
Antonio → Montse, Eva, Judith	Judith → Kepa, Antonio, Jordi

En **rojo** parejas inestables

a=Jordi y Montse, **b**=Kepa y Eva, **c**=Antonio y Judith

//d=Jordi y Montse, e=Kepa y Judith, f=Antonio y Eva

Algoritmo de Gale-Shapley(1)

David Gale y Lloyd Shapley demostraron que siempre es posible encontrar un conjunto de parejas estables:

- Se forma una lista con todas las mujeres sin pareja, inicialmente todas.
- La primera mujer sale de la lista y trata de emparejarse con su favorito
- Si ese hombre no tiene pareja, ambos se unen y se pasa a la siguiente mujer
- Si el hombre ya tiene pareja, examina si prefiere a esta o a su pretendiente y forma pareja con la mujer que más le guste de las dos. Si la mujer que le pretendía no es elegida pasa a su siguiente favorito
- El procedimiento se repite hasta que no hay ninguna mujer en la lista sin parejas.

Algoritmo de Gale-Shapley(2)

- Su generalización y variaciones se usan entre otras cosas:
 - Asignación de estudiantes residentes a hospitales
 - Asignación de órganos a enfermos en trasplantes
 - Repartir estudiantes en colegios mayores
 - Decidir qué *cluster* se utilizará para descargar contenido de un CDN
- Inconveniente
 - Coste temporal: Muchas iteraciones

Software desarrollado

Para averiguar si un clasificador puede sustituir al algoritmo de Gale-Shapley se ha creado un software que:

- Crea una serie de agentes de dos clases, elementosX y elementosY
- Estos agentes nacen con unas características y preferencias por esas características
- Cada agente se comunica con los del otro grupo, obtiene sus características y los ordena en una lista de favoritos
- Un agente emparejador forma un conjunto de parejas estables y no estables y los guarda en un archivo csv
- Un agente clasificador utiliza esos archivos csv como conjuntos de entrenamiento o prueba
- El usuario puede analizar el resultado de la clasificación

Opciones del programa

- **Un entero:** Número de iteraciones en la creación de parejas para obtener conjuntos de datos mayores
- **“completo”:**
 - GaleShapley (Guarda en el conjunto de entrenamiento como parejas inestables solo parejas formadas que se rompen)
 - GaleShapleyCompleto (Guarda como parejas inestables también las que son desechadas y no se llegan a formar)
- **“moduladas”:** Los agentes asignan a unas determinadas características más valor que a otras
- El sistema no permite elegir el algoritmo de clasificación, que siempre será Logistic (pruebas preliminares mostraron que se trataba del más prometedor)

Archivos de datos csv

El agente emparejador guarda en cada línea del archivo:

- **Pareja** Qué elementos forman la pareja
- **Características y preferencias** elementoX
- **Características y preferencias** elementoY
- **Clase** estable o no estable

Cada línea será una instancia para el agente clasificador

El nombre del archivo guarda los datos de configuración

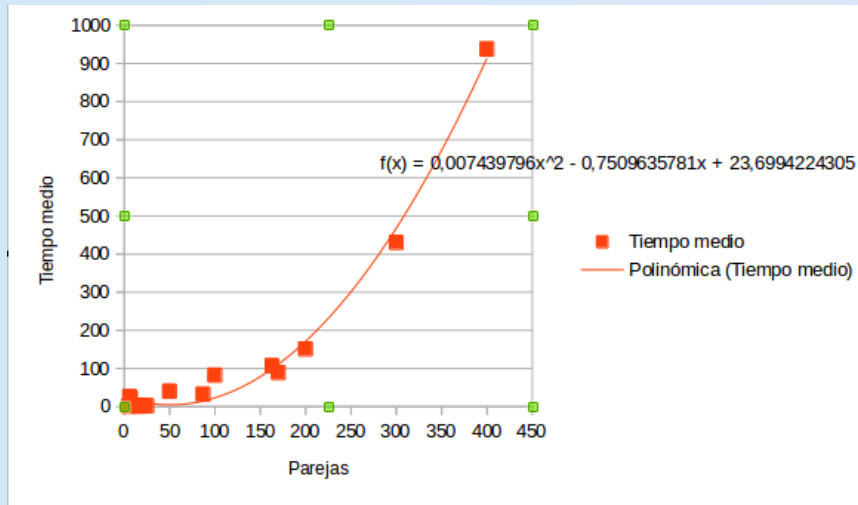
GaleShapleyCompleto_parejas_300_Iteraciones_1_2017_12_10_15_12_24modulacion:true.csv

Archivos de resultados csv

El agente clasificador guarda en cada línea:

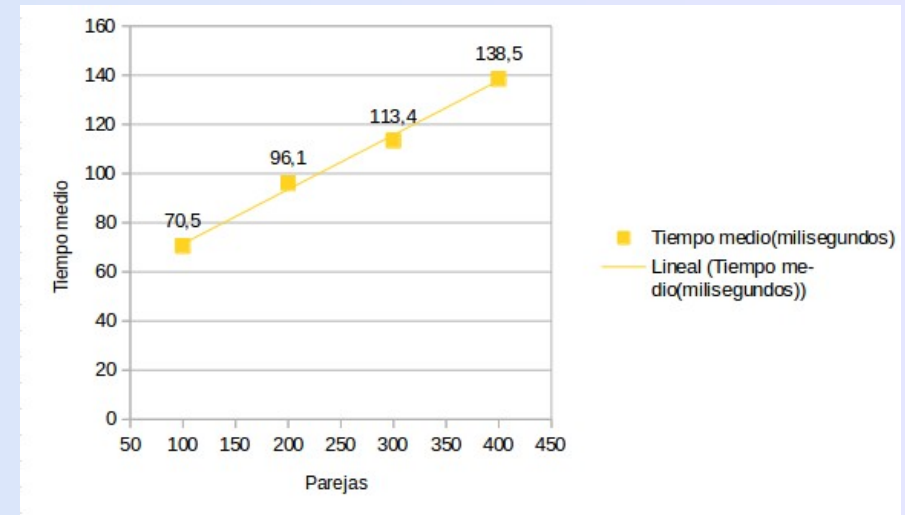
- Nombre de los archivos de entrenamiento y prueba
- Parámetros de ejecución de ambos conjuntos
Número de parejas, iteraciones, completo, modulación
- Resultados de la evaluación del clasificador
 - Tiempo empleado
 - Tasas de True Positive de cada clase
 - Precisiones de cada clase
 - F-Measures de cada clase
 - Estadístico Kappa

Resultados tiempo



Tiempos promedio de algoritmo de Gale-Shapley

Crecimiento polinómico



Tiempos promedio de evaluación clasificadores

Crecimiento lineal

El tiempo que emplean los clasificadores en formar parejas estables es mucho menor que el algoritmo de Gale-Shapley

Tasa True Positive y Precisión(Estables)

- Tasa True Positive:

Indica el porcentaje de parejas estables que realmente ha encontrado el clasificador.

Una tasa de 0.5 indicaría un 50%

- Precisión:

Indica el porcentaje de parejas señaladas como estables son realmente estables

Mejores clasificadores para números grandes de parejas

Parejas	Iteraciones	Completo	Modulacion	Parejas	Iteraciones	Completo	Modulacion	Tasa TP (Estables)	Precision Estables
200	5	NoCompleto	true	200	1	Completo	true	0,48	0,2
200	1	NoCompleto	true	200	1	Completo	true	0,44	0,2
200	3	NoCompleto	false	200	1	Completo	false	0,43	0,17
200	3	NoCompleto	true	200	1	Completo	true	0,41	0,21
200	5	NoCompleto	false	200	1	Completo	false	0,37	0,23
400	3	NoCompleto	false	400	1	Completo	false	0,36	0,17
200	1	NoCompleto	false	200	1	Completo	false	0,35	0,2
400	5	NoCompleto	false	400	1	Completo	false	0,32	0,18
400	1	NoCompleto	true	400	1	Completo	true	0,32	0,16
400	1	NoCompleto	false	400	1	Completo	false	0,3	0,16
300	1	NoCompleto	false	300	1	Completo	false	0,26	0,18
400	5	NoCompleto	true	400	1	Completo	true	0,25	0,15
300	1	NoCompleto	true	300	1	Completo	true	0,24	0,21
400	3	NoCompleto	true	400	1	Completo	true	0,24	0,18
300	5	NoCompleto	false	300	1	Completo	false	0,23	0,21

Entrenamiento

Conjunto de test

Resultados

- Las Tasas de True Positive obtenidas por los clasificadores son muy bajas
- Las precisiones obtenidas también son muy pobres
- No hay una correlación clara entre los parámetros de ejecución y la calidad de un clasificador

No es posible sustituir el algoritmo de Gale-Shapley por un sistema clasificador

Conclusiones

- **Objetivos del trabajo cumplidos**
 - Un clasificador no puede sustituir al algoritmo de Gale-Shapley
 - Objetivos personales
 - Aprendizaje de SMA
 - Desarrollo completo de un software utilizando diferentes técnicas
- **Posibles líneas de ampliación**
 - Uso de otros algoritmos de clasificación (Improbable resultado positivo)

Muchas gracias

**Quedo a su disposición para
cualquier pregunta**

<mailto:ahuesoa@uoc.edu>