

Organización de Asignaturas en una Facultad mediante Algoritmos Genéticos

Fernando Nellmédin y Mariano Peyregne

Universidad Nacional del Litoral, Facultad de Ingeniería y Ciencias Hídricas

Resumen—El problema de acomodar los horarios de una facultad, llamado *TimeTabling Problem* en inglés, es una difícil tarea que se debe realizar en cualquier facultad o universidad. Se tienen un conjunto de recursos limitados tales como la cantidad de profesores, alumnos o aulas, y se los debe asignar a cada una de las asignaturas que se cursan en la institución de manera de minimizar el impacto negativo tanto en los alumnos como en los docentes. Se quieren evitar situaciones tales como que se cursen simultáneamente materias del mismo año o de años adyacentes, ya que con frecuencia hay estudiantes que pueden asistir a asignaturas de distintos años. En este paper abordamos el *Timetabling Problem* utilizando un Algoritmo Genético y modelando las restricciones en una simple función de fitness. Si bien no consideramos todas las restricciones que participan en este problema, encontramos resultados satisfactorios con respecto al programa semanal obtenido. Primero presentamos el modelado del problema, luego los operadores definidos para el algoritmo genético y el fitness. Por último se presentan los resultados obtenidos y algunos ejemplos.

Palabras clave—*timetabling problem*, restricciones, algoritmo genético, optimización

I. INTRODUCCIÓN

La distribución de los horarios en que se dictan las asignaturas de la facultad se debe actualizar todos los años e implica un trabajo iterativo hasta que se cumplan todas las restricciones, o al menos se minimicen la cantidad de conflictos. Entre los requisitos a cumplir se encuentran los inherentes a los planes de las carreras, tales como que no se asignen materias del mismo año al mismo horario, o incluso de años consecutivos, ya que muchos alumnos asisten a asignaturas de distintos años por problemas de correlatividades. Asimismo, otras limitaciones surgen por la disponibilidad de aulas (y sus correspondientes equipamientos y límite de alumnos) y los docentes asignados a cada clase. Sobre este último, con frecuencia surge que cada profesor impone una restricción del tipo: "sólo de mañana" o "los lunes y martes únicamente". De la definición de todas estas limitaciones, vemos que encontrar una distribución de clases durante la semana se torna una tarea casi imposible para cualquier ser humano.

En este trabajo, nos enfocamos en solucionar el *Timetabling Problem* (TTP) para un conjunto limitado de restricciones. Además, simplificamos la definición del problema. Especializamos el trabajo en nuestra propia carrera: Ingeniería en Informática. Por lo tanto, modelamos las asignaturas de los 5 años de la misma. También se resumieron las clases de cada materia a 3: una de teoría de dos horas, y dos de práctica de tres horas cada una (a las que los alumnos deben acudir a ambas). En total hay 20 materias que se dictan en un semestre, de 3 clases cada una, lo que suma un total de 60 clases a ubicar. Un ejemplo para las primeras asignaturas se ve en la Tabla I. Luego se estudiará el caso más general.

TABLA I
ASIGNATURAS MODELADAS COMO CLASES

Asignatura	Clase	ID
Cálculo II	Cálculo II (Teoría, 2 horas)	1
Cálculo II	Cálculo II (Práctica 1, 3 horas)	2
Cálculo II	Cálculo II (Práctica 2, 3 horas)	3
Física I	Física I (Teoría, 2 horas)	4
Física I	Física I (Práctica 1, 3 horas)	5
Física I	Física I (Práctica 2, 3 horas)	6

No se tuvo en consideración la distribución de los profesores ni sus restricciones añadidas, por lo que se asume que cada docente tiene a su cargo una sola clase y puede darla en cualquier franja horaria. En nuestra institución, como en muchas, se dicta más de una carrera. En lugar de modelar la distribución de clases y aulas para todas las carreras, se tomó un camino alternativo: se limitarán las aulas ocupadas por nuestra carrera, de manera que sólo se puede ocupar determinado número de aulas en simultáneo.

Por el lado de las restricciones, elegimos las siguientes a satisfacer:

- *Limitación de aulas*: No se pueden utilizar más de 4 aulas al mismo tiempo
- *Asignaturas en un mismo día*: No se desea que la misma asignatura, aún de clases distintas, se dicte dos veces en el mismo día.
- *Solapamiento en años adyacentes*: No se deben superponer las clases de asignaturas de un año, con las clases del año anterior o siguiente, para dar la posibilidad a los alumnos de cursar ambos semestres.
- *Superposición en un mismo año*: No es posible que se dicten dos clases del mismo año en el mismo momento, ya que imposibilita a los estudiantes a hacerse presente en ambas clases.

Con respecto a la distribución de horarios, se divide cada día, de lunes a viernes, en 5 módulos o franjas horarias. A cada módulo se le asigna un ID único, que servirá para identificarlo en el algoritmo genético. En la Fig. 1 se aprecian los módulos y el ID asignado a cada uno. El uso de estos identificadores se explicará en la próxima sección.

Horario	Lunes	Martes	Miércoles	Jueves	Viernes
8 a 10	0	5	10	15	20
10 a 13	1	6	11	16	21
14 a 16	2	7	12	17	22
16 a 18	3	8	13	18	23
18 a 21	4	9	14	19	24

Fig. 1. Módulos definidos para la semana

II. ALGORITMO GENÉTICO

Para la solución del problema, se utilizó un AG en el que se utilizó un fenotipo como se muestra en la Fig. 2. De acuerdo a los ID definidos para cada clase en la Tabla I, se le asigna una posición específica en el fenotipo a cada una. Así, la clase con identificador igual a 1, será la primera en el fenotipo. Lo mismo se aplica a ID = 2, que estará ubicada segunda, y lo mismo con el resto. En cada posición del fenotipo, se almacena un número que referencia en qué bloque está situada. Siguiendo con el ejemplo de la Fig. 2, se aprecia que la clase 1 se dictará en el módulo 12, que al consultar la Fig. 1 vemos que corresponde a los miércoles a las 14 horas. A partir de la representación fenotípica, se codifica en un genotipo binario de 300 bits de extensión, ya que se toman 5 bits para cada elemento del fenotipo (que está en el rango 0-24).

12	13	16	1	3	...	20
----	----	----	---	---	-----	----

Fig. 2. Representación fenotípica

A. Inicialización

Se parte de una población inicial aleatoria. Para generarla, se toman las clases de a una y se le asigna un bloque aleatorio en el rango 0-24. De esta forma, se comenzará con muchos individuos con una evaluación de fitness pobre, ya que no se espera que cumplan ninguna restricción al inicio, pero todos serán válidos en el sentido de que no se asignarán bloques fuera del rango 0-24. Esta propiedad se mantendrá durante todo el algoritmo.

B. Selección de individuos para la siguiente generación

La población siempre se mantiene en el mismo tamaño. En cada generación, se deben elegir a los individuos que formarán parte de la generación siguiente. Se utilizó el método de selección por ventanas. A los individuos elegidos se les aplicará, si corresponde, las cruza y mutaciones que se describen en la próxima sección. Además, a través de elitismo, siempre se pasa el mejor individuo (según su fitness) a la próxima generación. Por último, se elige un término de brecha generacional del 10% del tamaño de la población. Por lo tanto, sin importar el fitness, se elige aleatoriamente cierta cantidad de individuos que serán integrantes del próximo conjunto. De esta manera, nos aseguramos que siempre haya diversidad en el conjunto de búsqueda.

C. Operadores genéticos

Se definieron dos operadores de mutación y uno de cruza. Estos trabajan en representación genotípica, pero cuidan de realizarse en posiciones que se corresponden con el fenotipo, de manera de no generar individuos inválidos luego de la operación. Cada uno de estos operadores tiene una probabilidad de aplicarse definida al inicio, por lo que puede pasar que ambas mutaciones se apliquen y la cruza no, o viceversa o cualquier otra combinación.

C.1. Cruza: Se utiliza una cruza de dos puntos. Se cuenta con dos padres y se generan aleatoriamente dos posiciones en el genotipo de ambos. Luego se intercambia el contenido genético entre éstos. Las posiciones elegidas se corresponden con los límites de cada elemento del fenotipo, de manera de no generar individuos con bloques de horarios inválidos.

C.2. Mutación por movimiento: Se genera una posición en el fenotipo al azar, que identificará a un bloque donde está asignada cierta asignatura. A continuación se buscará un módulo que se encuentre libre y que sea del mismo tamaño (2 o 3 horas según corresponda). Si se encuentra un espacio libre (después de un número limitado de iteraciones), se moverá la clase al nuevo bloque. En la Fig. 3 se ve un ejemplo.

Horario	Lunes	Martes	Miércoles	Jueves	Viernes
8 a 10	Mat	Química			
10 a 13		POO	APS	Mat 2	PDS
14 a 16			IComp		
16 a 18					
18 a 21					

Horario	Lunes	Martes	Miércoles	Jueves	Viernes
8 a 10	Mat	Química		IComp	
10 a 13		POO	APS	Mat 2	PDS
14 a 16					
16 a 18					
18 a 21					

Fig. 3. Mutación por movimiento

C.3. Mutación por permutación: Se toma una posición aleatoria del fenotipo. A partir de ésta, y dependiendo de si es de 2 o 3 horas, se genera una nueva posición aleatoria del mismo horario. Para continuar, se intercambian las posiciones de ambas, de manera de que donde se ubicaba la primera, ahora se ubique la segunda y viceversa. En la Fig. 4 mostramos un ejemplo ilustrativo.

Horario	Lunes	Martes	Miércoles	Jueves	Viernes
8 a 10	Mat	Química			
10 a 13		POO	APS	Mat 2	PDS
14 a 16			IComp		
16 a 18					
18 a 21					

Horario	Lunes	Martes	Miércoles	Jueves	Viernes
8 a 10	Mat	IComp			
10 a 13		POO	APS	Mat 2	PDS
14 a 16			Química		
16 a 18					
18 a 21					

Fig. 4. Mutación por permutación

D. Función de fitness

La función de fitness que evalúa la calidad de la solución representada por cada individuo, debe tener en cuenta qué tanto cumple con las restricciones definidas. Para ello, el fitness está formado por 4 términos, uno por cada requisito a cumplir. Se probaron dos tipos de funciones de fitness, una aditiva y otra multiplicativa. Primero se expondrán cada uno de los términos y luego la relación entre éstos.

D.1. Sobrepasso de Aulas: Cuenta la cantidad de veces que se supera la capacidad de aulas definida. Si se cuenta con 4 aulas, contará para cada bloque horario, la cantidad de clases que se dictan simultáneamente. Si se sobrepasa el límite de 4, se sumará 1 en este contador.

D.2. Repeticiones de Asignaturas: Cuenta las veces que se da más de una vez cierta asignatura en el mismo día. Si el individuo establece, por ejemplo, que la teoría y la práctica 1 se dan en el mismo día, sumará 1 en este contador.

D.3. Solapamiento de años adyacentes: Se cuentan la cantidad de oportunidades en las que clase de años adyacentes se dictan en el mismo módulo. Por ejemplo, este contador sumará 1 cada vez que el individuo represente que un curso de segundo año se solape con cualquier clase de primer o tercer año.

D.4. Superposición de clases: Se aumentará el contador de esta restricción cuando dos clases del mismo año se asignen al mismo bloque horario, ya que significará que un alumno debería estar al mismo tiempo en dos clases distintas.

III. RESULTADOS

IV. CONCLUSIONES

A. Trabajos futuros

V. REFERENCIAS

La extensión de cada trabajo no podrá exceder *estrictamente* las 4 páginas y se recomienda que tampoco posea menos de 3 páginas. Por favor, no cambie la tipografía, ni los espacios entre el texto, ni otras medidas definidas en los estilos para ajustarse a esta norma. En esta oportunidad se solicita a los autores que para la revisión inicial presente el trabajo completo (3 o 4 páginas) en el formato final pero sin los datos de autor y por separado un archivo de texto con el título, los datos de los autores, el resumen y las palabras clave.

La estructura general que se espera para este artículo abarca generalmente secciones como: introducción, materiales, métodos, resultados, discusión, conclusiones, trabajos futuros, agradecimientos y referencias. Estos títulos pueden combinarse de a dos en una misma sección y los títulos trabajos futuros y agradecimientos son totalmente optativos. Es común que la sección de métodos lleve otro título más relacionado con el aporte original del artículo, pero las restantes secciones se presentan con los títulos antes listados. En la sección de discusión no debe faltar un análisis crítico de los métodos y resultados obtenidos. Si existieran demostraciones u otros desarrollos matemáticos extensos, se recomienda agruparlos en apéndices antes de las referencias bibliográficas.

A continuación se darán más detalles acerca de las secciones del documento y los formatos para insertar los distintos tipos de objetos, como ecuaciones, figuras, etc.

VI. FORMATO PARA LOS OBJETOS INSERTADOS

Observe que el título de cada sección ya incluye la numeración automática y el espacio hacia el último párrafo de la sección anterior. Por lo tanto, para iniciar una nueva sección se recomienda comenzar copiando y pegando el título desde otra sección en este mismo documento. De forma similar, también se puede observar que el formato para los párrafos ya incluye una pequeña indentación automática

en la primera línea y no existe un espacio extra para la separación entre párrafos.

A. Las ecuaciones

Las ecuaciones menores o definiciones de variables pueden insertarse directamente en la línea del párrafo, por ejemplo, considérese que se desea definir una historia $\mathbf{h}_i^n = w_{i-1}, w_{i-2}, \dots, w_{i-n+1}$ asociada un símbolo w_i . Observe que una manera sencilla de asegurar la uniformidad en el estilo de las ecuaciones es insertar siempre una ecuación aún cuando se podría escribir directamente como texto y aplicar formato en cursiva o negrita.

Para insertar ecuaciones más complejas se recomienda utilizar un formato de párrafo aparte, con el estilo correspondiente:

$$\hat{P}_I(w_i|\mathbf{h}_i^k) = \sum_{j=0}^{k-1} \lambda_j \hat{P}(w_i|\mathbf{h}_i^j) \quad (1)$$

En este estilo de ecuación se ha fijado dos tabulaciones, la primera centra la ecuación en la columna y la segunda justifica a la derecha el número de la ecuación. Para hacer referencia a esta ecuación desde el texto se menciona, por ejemplo, en (1) se puede ver la estimación de la probabilidad de una historia a partir de la simple combinación lineal de historias de orden inferior.

B. Las figuras

Cuando inserte figuras no las deje como objetos flotantes sobre el texto y no incluya dentro de ellas al epígrafe. El epígrafe se coloca abajo de las figuras y posee un estilo propio que básicamente consiste en fuente de 8 pt con párrafo justificado cuando se trate de más de una línea y centrado para una única línea de texto (ver Fig. 1). Si en la figura utiliza ejes cartesianos, recuerde siempre indicar en la misma a que corresponde cada eje (etiquetas). Para hacer referencia a una figura se debe utilizar la forma abreviada Fig. seguida del número de la figura, salvo cuando esté al comienzo del párrafo, caso en que se deberá utilizar la palabra completa.

En la Fig. 2 se puede ver otro tipo de figura donde se destacan cuatro curvas. No incluya colores en las gráficas, preferentemente utilice distintos tipos de líneas. Los gráficos vectorizados brindan una mejor calidad electrónica y de impresión. Por lo tanto, inserte todas las gráficas con algún formato vectorizado (Corel, Visio, XFig, PostScript, Metarchivo mejorado, etc) o bien, si se tratase de una fotografía o imagen más compleja utilice formatos raster sin compresión (por ejemplo BMP) o con compresión sin pérdida de información (se pueden configurar los formatos JPG, PNG, TIF, GIF, etc).

En general, se recomienda que las figuras estén al principio o al pie de una columna. Se debe cuidar que el tamaño del texto dentro de las figuras no sea menor a 7 pt.

C. Las tablas

Es preferible que las tablas se diseñen a partir del mismo editor de textos pero también pueden consistir en una gráfica en algún formato vectorizado. El epígrafe de las tablas es marcadamente diferente del de las figuras: se coloca por arriba de la tabla, con fuente de tamaño 8 pt en versales y párrafo centrado.

Al igual que en las figuras, es preferible que las tablas se encuentren al principio o al pie de una columna. El tamaño del texto dentro de las tablas no debería ser inferior a 7 pt ni mayor a 10 pt.

Un ejemplo de este estilo puede verse a continuación en la Tabla I.

TABLA II

RESULTADOS FINALES Y REDUCCIÓN RELATIVA DE LOS ERRORES
(PROMEDIOS SOBRE 10 PARTICIONES DE ENTRENAMIENTO Y PRUEBA).

Errores de reconocimiento	SER %	WER %	WAER %	Reducción %WER
Referencia	38.30	7.54	8.53	—
HMM-PASS	30.55	5.36	6.67	28.91
T-PASS	25.50	4.76	5.70	36.87

D. Las citas bibliográficas

Las citas bibliográficas se realizan entre corchetes, por ejemplo [1]. Cuando se hacen citas múltiples utilice la coma para dos citas [2], [3] o bien la notación de rangos de citas [2]-[5]. No utilice términos particulares antes de la cita, como en la “referencia [2]” o en “Ref. [4]”. El estilo general para las referencias bibliográficas se muestra con varios ejemplos en la sección correspondiente al final de este documento. Observe estrictamente el estilo propuesto: la utilización de tipografía, las mayúsculas, la forma de nombrar a los autores, los datos requeridos para libros, revistas y congresos, etc.

La sección de referencias posee un estilo particular para el párrafo y la numeración. Éstas se deben presentar preferentemente por orden de aparición en el texto pero también se aceptarán por orden alfabético del apellido del primer autor.

E. Otras recomendaciones generales

Defina adecuadamente cada uno de los acrónimos la primera vez que aparece en el texto (salvo en el resumen), por ejemplo, relación de grandes masas (RGM). Luego utilice siempre el acrónimo en lugar del término completo.

Recuerde definir cada uno de los símbolos que aparecen en las ecuaciones y aclarar la notación cuando se utilizan operadores matemáticos especiales o poco comunes.

Observe la utilización de mayúsculas, como regla general se coloca mayúscula en la primer letra de la primer palabra de cada frase y los nombres propios, tanto en los títulos como en el texto en general.

VII. FORMATO ELECTRÓNICO DE ENVÍO

Hasta la fecha indicada en el sitio web del congreso se recibirán: el artículo completo de 4 páginas en formato PDF y el título, resumen, palabras clave y datos de los autores en formato de texto ASCII puro.

Dado que el proceso de revisión se realizará por el sistema de doble ciego, solicitamos que *no* se incluyan los nombres y datos de autores en el artículo completo enviado para la revisión. En lugar de estos datos deje simplemente las leyendas bajo el título tal como figuran en este ejemplo. Estos datos deberán incluirse en el archivo de texto y en la versión final de documento completo (ya revisado, aprobado y listo para imprimir).

A continuación se detalla como generar cada uno de estos archivos.

A. Artículo completo para la revisión

El formato en que se reciben los artículos completos es PDF hasta la versión 1.4. Para los usuarios de \LaTeX se recomienda utilizar directamente el programa `pdflatex`. Si no contara con este programa en alguna distribución de \LaTeX , también es posible utilizar la secuencia `latex`, `dvips` y `ps2pdf`.

Para los usuarios de Word que no conozcan un método para convertir sus documentos a PDF, le damos a continuación una serie de instrucciones que le facilitarán esta tarea.

En primer lugar deberá descargar e instalar los siguientes programas gratuitos (en este mismo orden):

1. Impresora PostScript¹: se instala como una impresora y le permitirá imprimir el documento en formato PS. Durante la instalación elija FILE como puerto para la impresora.
2. Intérprete PostScript²: recomendamos en este caso el intérprete GhostScript, que se encuentra disponible gratuitamente para diferentes sistemas operativos.
3. Visualizador PostScript³: GSView utiliza al intérprete anterior para previsualizar archivos PS y para convertirlos en PDF.

Para la conversión PDF se utiliza el formato intermedio PS. En primer lugar se imprime el documento en la impresora PostScript. El archivo quedará con extensión `.prn` y es necesario cambiarla manualmente a `.ps`. Una vez obtenido el archivo PS se debe abrir con GSView y seleccionar la opción para “convertir...” del menú de archivos. Mediante la opción `pdfwrite` y utilizando la máxima resolución disponible, podrá convertir finalmente todas las páginas de su PS a PDF.

B. Resumen y datos de autores en texto ASCII

En un archivo de texto estándar (`.txt`) y sin caracteres ni formatos especiales se deben incluir los siguientes datos en castellano y en inglés:

- Título del trabajo
- Nombre completo del 1er autor
- Institución a la que pertenece el 1er autor
- Datos del 1er autor (dirección, TE/FAX, email, etc)
- Nombre completo y datos de los coautores
- Resumen (el mismo del artículo completo)
- Palabras clave

C. Versión final del artículo

Una vez aprobado para la publicación e implementadas las correcciones sugeridas por los revisores, se deberá enviar el artículo final en formato PDF. Este archivo deberá tener como nombre el mismo código entregado durante la revisión (COD-PAIS_NRO) y `.pdf` como extensión.

La versión final deberá incluir, debajo del título, los nombres de los autores y la información de la institución a la que pertenecen (tal como se indica en este ejemplo).

¹<http://www.adobe.com/products/printerdrivers/main.html>

²<http://www.cs.wisc.edu/~ghost/doc/AFPL/get704.htm>

³<http://www.cs.wisc.edu/~ghost/gsview/>

VIII. CONCLUSIONES

En las conclusiones debería presentarse una revisión de los puntos clave del artículo con especial énfasis en el análisis y discusión de los resultados que se realizó en las secciones anteriores y en las aplicaciones o ampliaciones de éstos. No debería reproducirse el resumen en esta sección.

IX. APÉNDICES

En algunas situaciones conviene incluir una sección de apéndices con sus correspondientes subsecciones.

A. *Demostraciones*

Cuando la extensión y complejidad de las demostraciones lo justifique en pos de no distraer al lector presentando en el texto solamente los resultados finales.

B. *Algoritmos*

Cuando sus extensiones lo justifiquen y no sean parte central del trabajo.

C. *Detalles técnicos*

Tablas con datos técnicos o mediciones accesorias que se utilizaron en el trabajo.

AGRADECIMIENTOS

Si los hubiere, a quienes corresponda.