

Micro-Proyecto 1: Enseñanza

Alejandro Noriega Soto

Juan David Correa Arango

Santiago Salazar Ramirez

Universidad Nacional de Colombia sede Medellín

1. Problemas por resolver:	3
2. Justificación:	3
3. Descripción del FIS	5
Reglas de inferencia difusa:	6
Capturas del despliegue del FIS_Examen:	10
Casos de estudio del FIS_Examen:	11
Caso de estudio 1:	11
Caso de estudio 2:	12
4. Descripción del SE	14
Descripción:	14
Consultas al SE_Recomendacion_Horas:	16
Consulta 1:	16
Consulta 2:	17

1. Problemas por resolver:

Tema asignado: Enseñanza

En el campo de la educación existe una problemática asociada con la efectividad de los métodos de evaluación sobre las temáticas aprendidas de un curso, donde uno de los defectos de los estudiantes es no conocer la preparación necesaria para un buen desarrollo de sus evaluaciones, con lo cual podrá afectar de forma completa la experiencia académica. El problema concreto es el manejo del tiempo por parte de los estudiantes para la preparación a un examen.

La **función principal** del FIS es encontrar la relación entre el tiempo de duración de un examen, la cantidad de preguntas, el porcentaje de aprobación y el rendimiento general de los estudiantes con la dificultad que maneja el examen (dificultad y porcentaje de pérdida del examen). El SE (sistema experto) tomará la dificultad y porcentaje de pérdida para encontrar la cantidad de horas recomendadas de estudio, esto con el fin de poder hacer más llevadera la experiencia académica.

2. Justificación:

Justificaciones para utilizar lógica difusa en este sistema:

1. No existe un convenio para la definición de lo que es fácil o difícil, de lo que es pocas o muchas preguntas, poco o mucho tiempo. Al no existir una clara definición de esto (otra de las dificultades del lenguaje) se busca emplear una aproximación a la estimación con la lógica, por eso se usa lógica difusa.
2. Teniendo en cuenta los conocimientos básicos acerca de lógica (lógica proposicional), ¿cómo sería posible definir este sistema de forma efectiva con

el solo uso de lógica clásica? No darían resultados claros sobre lo que se está hablando, siendo engorroso el proceso de creación del sistema.

3. Ya luego con la definición de las variables lingüísticas se puede hablar en resultado numérico, haciendo más claro el proceso de transformación de variable lingüística a un valor tangible. Esa es una de las mayores ventajas del uso de la lógica difusa, que se usan valores comprensibles por el usuario (es preferible decir que la temperatura ambiente está a -2°C que decir que está helando afuera).

Nombre del sistema: Nombre del FIS: FIS_Examen.

Nombre del SE: SE_Recomendacion_Hora

Utilidad del SE: El objetivo de toda institución académica es prestar una educación de calidad a sus clientes (estudiantes), pero en algunas ocasiones existe un mal rendimiento en los estudiantes por falta de preparación para la vida académica. El SE tendrá la utilidad de prestar una **APROXIMACIÓN** (el tiempo de estudio de cada estudiante se basa más en atributos personales, la dificultad del examen solo le da una idea de cuánto se debería preparar) de tiempo mínimo de estudio, basado en la dificultad estimada del examen. Se buscaría entonces mejorar la productividad de los estudiantes, apostar por la permanencia de ellos y por el desarrollo académico adecuado.

3. Descripción del FIS

Variables lingüísticas de entrada con sus universos de discurso:

1. Tiempo de duración límite: 15 minutos a 240 minutos.
2. Cantidad de preguntas: 3 preguntas a 40 preguntas.
3. Aprobación: 0 a 100.
4. Rendimiento: 0 a 5.

Variables lingüísticas de salida con sus universos de discurso:

- Dificultad: 0 a 10.
- Pérdida: 0 a 100.

Conjuntos difusos y funciones de pertenencia para las variables lingüísticas de entrada:

- Tiempo de duración límite: muy poco(Función de decremento lineal: 15 minutos a 40 minutos), poco(Función trapezoidal: 40-70, sus puntos son: 40,50,60,70), medio(Función trapezoidal: 60-130, sus puntos son: 60,80,105,130), alto(Función trapezoidal: 115 a 180, sus puntos son: 115,130,160,180) y muy alto(Función de incremento lineal: 180 a 240 minutos).
- Cantidad de preguntas: muy poco (Función de decremento lineal: 3 a 10), intermedio(Función trapezoidal: 8-18, sus puntos son: 8,12,15,18), alto(Función trapezoidal: 15-30, sus puntos son: 15,19,24,30), demasiado(Función de incremento lineal: 28 a 40).
- Aprobación: pocos (Función de decremento lineal: 0 a 20), algunos (Función trapezoidal: 18-40, sus puntos son: 18,25,32,40), promedio (Función trapezoidal: 35-70, sus puntos son: 35,48,60,70), alto(Función de incremento lineal: 68 al 100).

- Rendimiento: muy bajo (Función de decremento lineal: 0 a 1.5), bajo (Función trapezoidal: 1.2-2.7, sus puntos son: 1.2,1.6,2.3,2.7), promedio (Función trapezoidal: 2.6 - 3.5, sus puntos son: 2.6,2.9,3.2,3.5), alto (Función trapezoidal: 3.4 a 4.5, sus puntos son: 3.4,3.7,4.1,4.5), muy alto (Función de incremento lineal: 4.4 a 5).

Conjuntos difusos y funciones de pertenencia para las variables lingüísticas de salida:

- Dificultad: muy fácil (Función de decremento lineal: 0 a 2.5), fácil (Función trapezoidal: 2.3 a 4.5, sus puntos son: 2.3,2.9,3.7,4.5), intermedio (Función trapezoidal: 4.2 a 7.5, sus puntos son: 4.2,5.2,6.3,7.5), complejo (Función de incremento lineal: 7.3 a 10).
- Pérdida: muy baja (Función de decremento lineal: 0 a 18), baja (Función trapezoidal: 16 a 35, sus puntos son: 16,22,28,35), regular (Función trapezoidal: 30 a 60, sus puntos son: 30,41,53,60), alta (Función de incremento lineal: 58 a 100).

Reglas de inferencia difusa:

- Tiempo de duración(poco) AND cantidad de preguntas (alto) -> Pérdida (alta).
- Aprobación(pocos) AND rendimiento (muy bajo) -> Pérdida(alta).
- Tiempo de duración(alto) AND cantidad de preguntas(intermedio) -> Pérdida (muy baja).
- Aprobación(promedio) AND rendimiento(promedio) -> Pérdida(regular).
- Tiempo de duración(poco) AND cantidad de preguntas (demasiado) -> Dificultad(complejo).

- Tiempo de duración(alto) AND cantidad de preguntas(intermedio) -> Dificultad(muy fácil).
- Aprobación(algunos) AND rendimiento(alto) -> Dificultad(intermedio).
- Cantidad de preguntas(muy poco) AND tiempo de duración(muy poco) AND aprobación(algunos) AND rendimiento (bajo) -> Dificultad(intermedia) AND pérdida(regular).
- Tiempo de duración (muy alto) AND cantidad de preguntas (alto) AND aprobación(promedio) -> Dificultad(intermedio).
- Tiempo de duración (medio) AND aprobación(alto) AND cantidad de preguntas(intermedio) -> Dificultad(muy fácil) AND pérdida (muy baja).
- Tiempo de duración (medio) AND cantidad de preguntas (intermedio) AND rendimiento (promedio) -> Dificultad(fácil) AND pérdida(baja).
- Tiempo de duración (medio) AND cantidad de preguntas (muy poco) AND rendimiento (muy alto) -> Dificultad (muy fácil) AND pérdida (muy baja)
- Aprobación (alto) -> Dificultad(fácil) AND pérdida(baja)
- Tiempo de duración(muy alto) AND aprobación (poco) -> Dificultad(complejo) AND pérdida(alta).
- Tiempo de duración (alto) AND rendimiento (bajo) -> Dificultad(intermedio)

Justificaciones de las reglas de inferencia difusa:

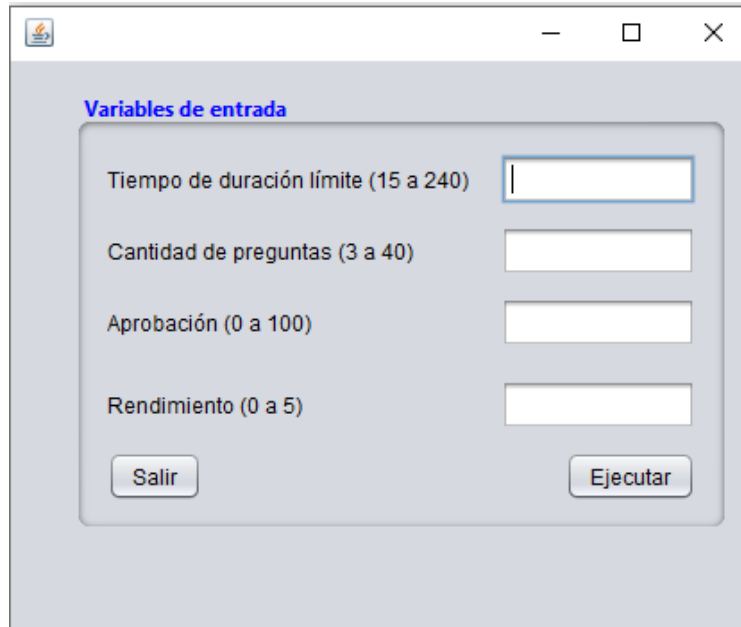
- Es de suponer que si el estudiante presenta una gran cantidad de preguntas en poco tiempo no podrá tener el suficiente tiempo para responder de forma

correcta, si es que siquiera alcanza a responder. La pérdida será por consiguiente alta.

- La aprobación y el rendimiento están estrechamente relacionados, siendo el rendimiento el promedio de los estudiantes, mientras que la aprobación se refiere al porcentaje de estudiantes que pasaron la materia. Entonces, si la aprobación es baja, significa que la pérdida es alta.
- Si el tiempo del estudiante para responder es alto y no tiene tantas preguntas para responder, significa que se tendrá un menor índice de pérdida.
- Esta regla se basa en la misma línea de la segunda regla, existe una relación entre aprobación y rendimiento. También hay que recalcar que existe una relación inversa entre la aprobación y la pérdida.
- Esta regla es una exageración de la primera, llevando al absurdo los tiempos y las situaciones.
- Relacionada a la anterior regla, pero su diferencia reside en el resultado (la dificultad ahora es muy fácil).
- Basado en el enunciado de que existe una relación entre aprobación y rendimiento, haciendo que cuando la aprobación sea media (algunos) y el rendimiento alto provoque una dificultad intermedia.
- Si el examen es "pequeño" (pocas preguntas y poco tiempo) y tiene un rendimiento bajo (pero su aprobación es intermedia) significa que la pérdida y dificultad es intermedia.
- Esta regla sigue un sentido similar a la anterior regla, pero se cambia a una dificultad y pérdida óptima (el examen es fácil).

- Las reglas 11 y 12 siguen el mismo funcionamiento, entonces no tiene sentido explicar otra vez la relación entre las variables
- Existe también una relación entre la dificultad y la pérdida, siendo una relación directa (entre mayor sea la dificultad mayor será la pérdida). Entonces, si la aprobación (que es opuesto de pérdida) es alto, significa que la dificultad y pérdida será fácil.
- El examen es largo y la aprobación es baja, el resultado esperado es que la dificultad del examen sea alta, junto a una pérdida alta.
- Esto mismo sucede con el rendimiento, se espera un resultado similar (dado que existe una relación entre la aprobación y el rendimiento).

Capturas del despliegue del FIS_Examen:



Variables de entrada

Tiempo de duración límite (15 a 240)

Cantidad de preguntas (3 a 40)

Aprobación (0 a 100)

Rendimiento (0 a 5)

La interfaz es sencilla en su aplicación. Se debe de rellenar los campos blancos con valores que estén dentro del rango (si son decimales se debe de usar el '.'), cuando ya estén listos los datos se oprime en ejecutar. En caso de que quiera salir de la aplicación debe de oprimir el botón de salir.

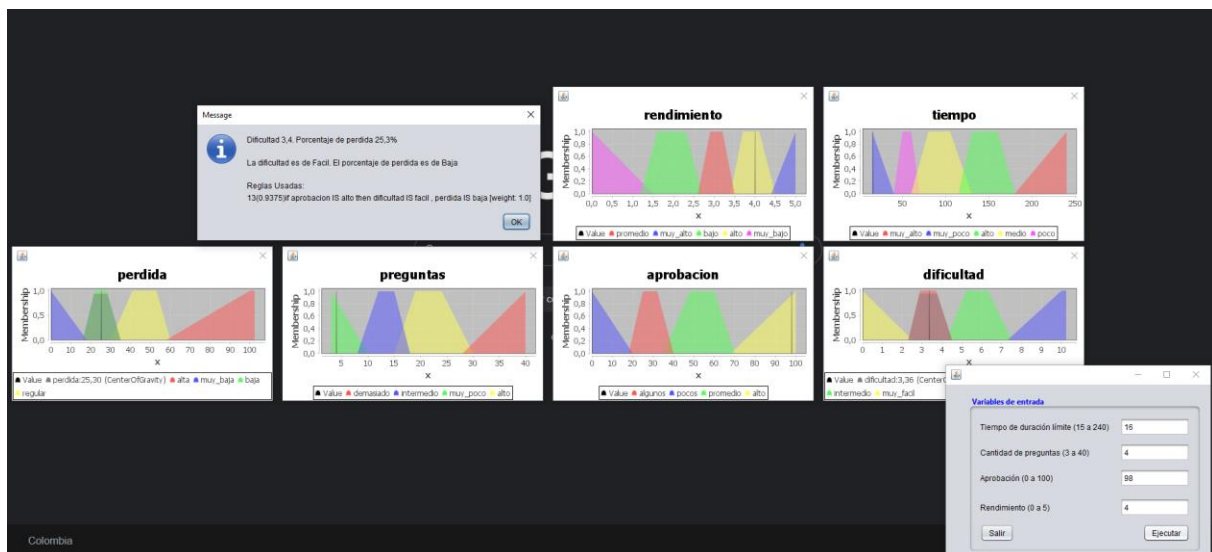


Cuando se ejecuta el programa muestra los resultados (Dificultad y pérdida), junto a los diagramas que muestra en qué conjunto difuso se encuentra el dato de la variable de entrada. También muestra el conjunto difuso en donde se encuentran las variables de salida y la regla que hace cumplir este resultado.

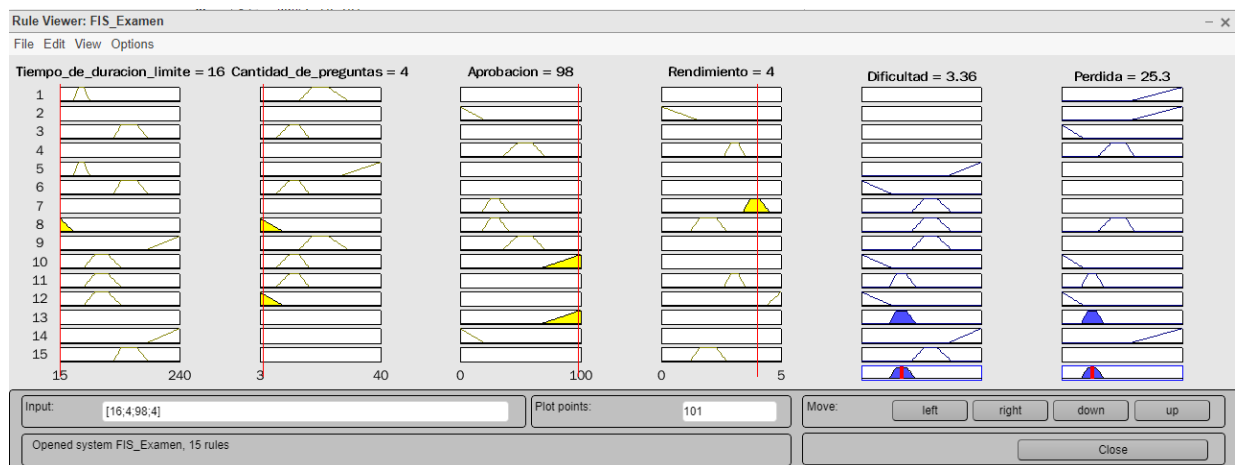
Casos de estudio del FIS_Examen:

Caso de estudio 1:

Java:



Matlab:



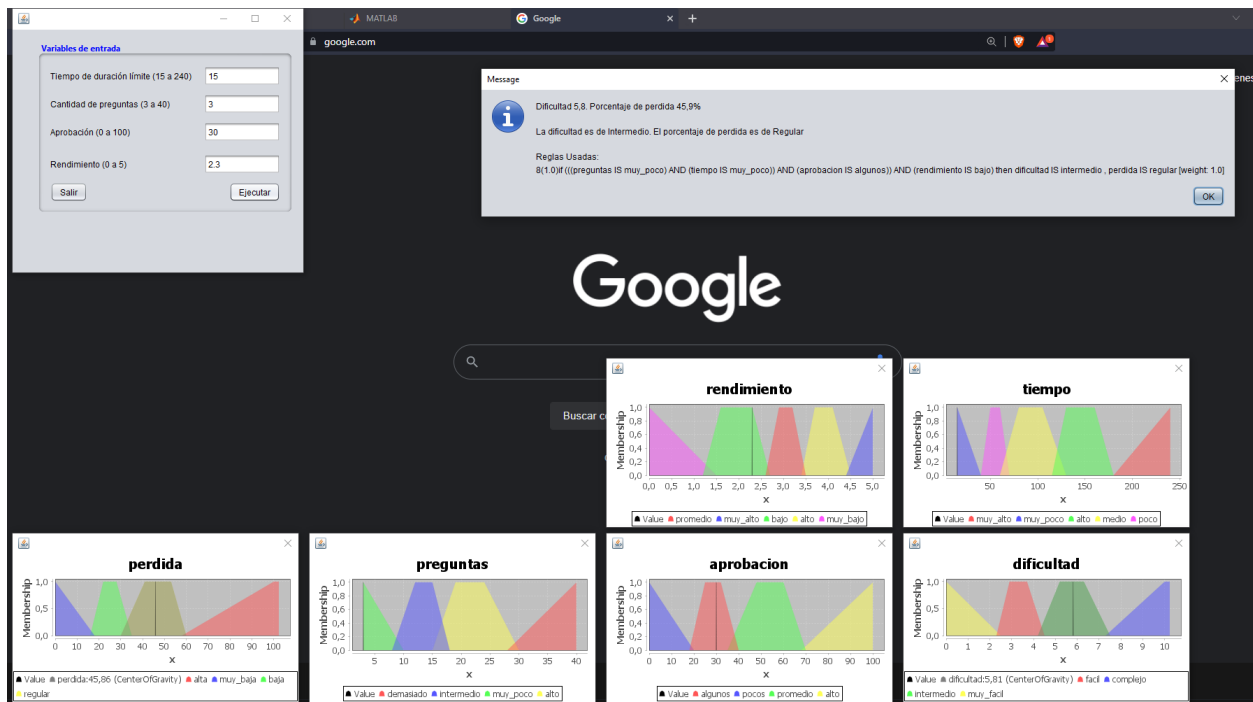
Datos de entrada:

- Tiempo de duración límite: 16
- Cantidad de preguntas: 4.
- Aprobación: 98.
- Rendimiento: 4.

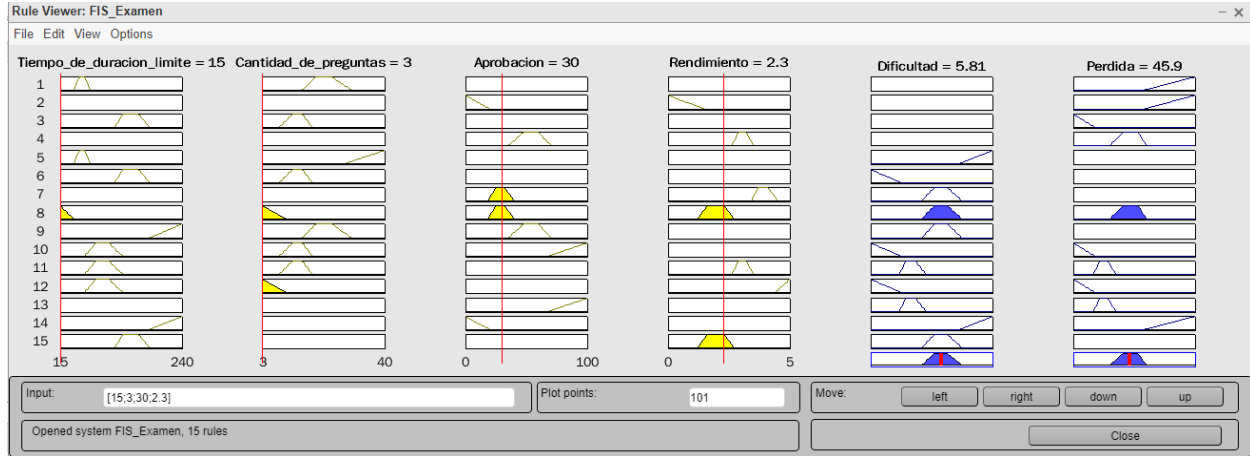
Explicación: Con los datos que se introdujeron se llega a que los conjuntos difusos de cada variable de entrada son: rendimiento es alto, tiempo de duración es muy poco, cantidad de preguntas son muy pocas, aprobación es alto. Como la aprobación es alta, entonces la dificultad es fácil y la pérdida es baja (se cumple esta regla). Los resultados que aparecen (3.4 en dificultad y 25.3% en pérdida) se basan en el centroide.

Caso de estudio 2:

Java:



Matlab:



Datos de entrada:

- Tiempo de duración límite: 15.
- Cantidad de preguntas: 3.
- Aprobación: 30.
- Rendimiento: 2.3.

Explicación: Con los datos que se introdujeron se llega a que los conjuntos difusos de cada variable de entrada son: rendimiento es bajo, tiempo es poco, preguntas es muy poco, aprobación es algunos. Como el rendimiento es bajo, tiempo es poco, preguntas es muy poco, aprobación es algunos; entonces la dificultad es intermedia y la pérdida es regular (se cumple esta regla). Los resultados que aparecen (5.8 en dificultad y 45.9%) se basan en el centroide.

4. Descripción del SE

Descripción:

El SE_Recomendación_Horas consta de un sistema de recomendación que le ayudará al usuario a determinar la cantidad de horas de estudio necesarias para aprobar satisfactoriamente el examen planteado desde los datos ingresados al FIS_Examen y, a su vez, desde los datos obtenidos a través del proceso de fuzzificación y defuzzificación los cuales son el porcentaje de pérdida y la dificultad estimada del examen.

Para dicho criterio, se tuvo en cuenta que a mayor dificultad y porcentaje de pérdida se requiere un mayor trabajo y, por consiguiente, en nuestro contexto, un mayor tiempo de estudio. Dichas estimaciones pueden verse reflejadas en la siguiente tabla de factores guía para realizar el SE_Recomendación_Horas:

Pérdida	Dificultad		Tiempo estudio mín	Tiempo estudio máx
muy baja	Muy fácil	<=====	0.5 horas	1.5 horas
baja	Muy fácil	<=====	1 hora	2 horas
regular	Muy fácil	<=====	1.5 horas	2.25 horas
alta	Muy fácil	<=====	1.75 horas	3 horas
muy baja	fácil	<=====	0.75 horas	2 horas
baja	fácil	<=====	1.5 horas	2.25 horas
regular	fácil	<=====	2 horas	3 horas
alta	fácil	<=====	2.5 horas	3.5 horas
muy baja	Intermedio	<=====	2.25 horas	3.25 horas
baja	Intermedio	<=====	2.5 horas	3.5 horas
regular	Intermedio	<=====	3 horas	4.5 horas
alta	Intermedio	<=====	3.5 horas	6 horas
muy baja	complejo	<=====	4.25 horas	6.5 horas
baja	complejo	<=====	4.75 horas	7 horas
regular	complejo	<=====	5.5 horas	8 horas
alta	complejo	<=====	6.5 horas	9.5 horas

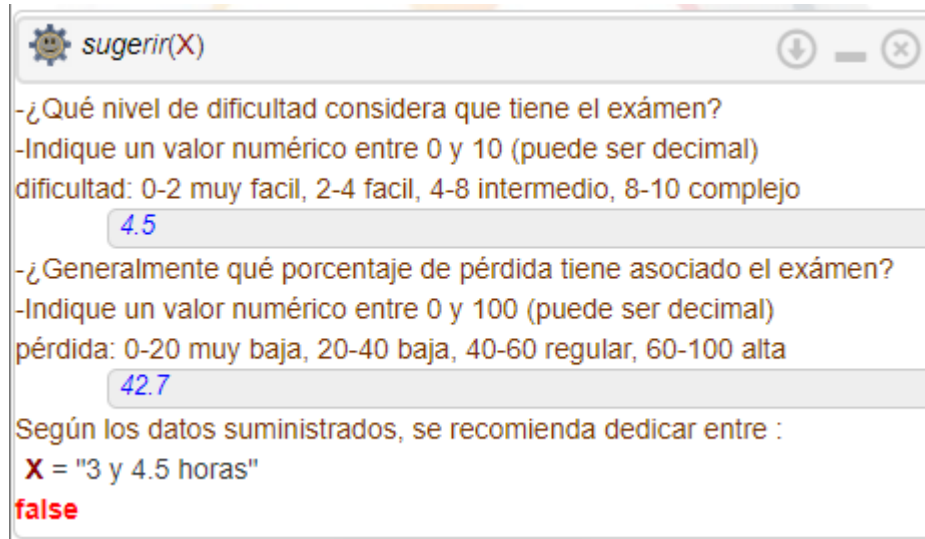
Es necesario recalcar de la tabla de factores anterior algunos aspectos:

- Tanto pérdida como dificultad poseen una relación similar con el tiempo de estudio estimado, pues a medida que alguna de las dos variables aumenta el tiempo de estudio también lo hace
- Sin embargo, en situaciones donde las variables alcanzan valores opuestos puede decirse que se atenúa el incremento en el tiempo de estudio estimado
- Si bien, para la estimación de los tiempos de estudio son importantes ambas variables, es necesario recalcar el peso mayor que tiene la dificultad sobre la pérdida al momento de ponderar los valores de la base de conocimientos del SE.
- Se prefirió generar un intervalo de tiempo de estudio en la recomendación de salida del SE para brindarle al usuario mayor comodidad al momento de generar una recomendación sobre algún examen, pues el tiempo de estudio, como unidad de medida del trabajo es bastante relativo, porque cada estudiante no asimila los conocimientos de la misma manera y con los intervalos de recomendación se busca atenuar dicha desviación
- Para generar las recomendaciones del SE_Recomendacion_Horas (valga la redundancia) se tuvo como referencia los siguientes valores como extrapolación del dominio asignado a las variables lingüísticas de salida del FIS_Examen:

- **Dificultad:** 0-2 *Muy Fácil*, 2-4 *Fácil*, 4-8 *Intermedio*, 8-10 *Complejo*
- **Pérdida:** 0-20 *Muy Baja*, 20-40 *Baja*, 40-60 *Regular*, 60-100 *Alta*

Consultas al SE_Recomendacion_Horas:

Consulta 1:



sugerir(X)

-¿Qué nivel de dificultad considera que tiene el examen?
-Indique un valor numérico entre 0 y 10 (puede ser decimal)
dificultad: 0-2 muy facil, 2-4 facil, 4-8 intermedio, 8-10 complejo
4.5

-¿Generalmente qué porcentaje de pérdida tiene asociado el examen?
-Indique un valor numérico entre 0 y 100 (puede ser decimal)
pérdida: 0-20 muy baja, 20-40 baja, 40-60 regular, 60-100 alta
42.7

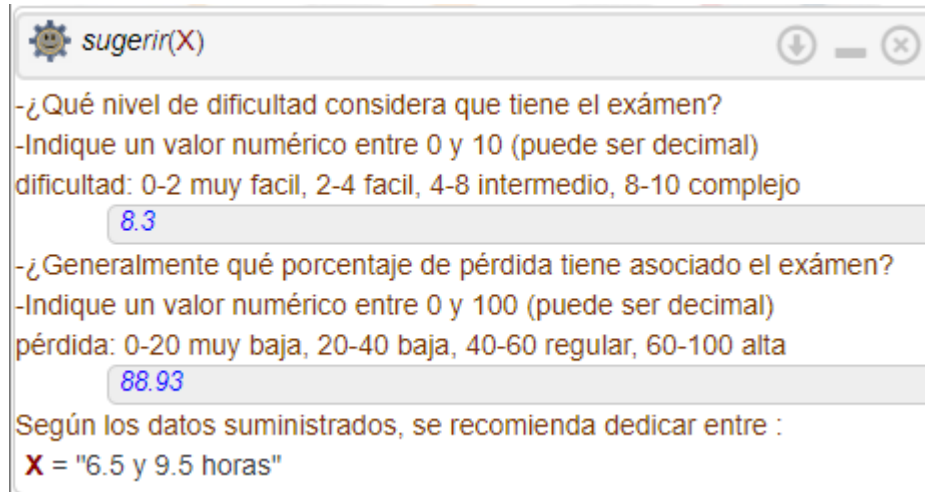
Según los datos suministrados, se recomienda dedicar entre :
X = "3 y 4.5 horas"
false




Posteriormente al ingreso de los valores numéricos que solicita el SE, se procede a realizar la búsqueda de la regla que entra en funcionamiento bajo los parámetros de entrada ya descritos. En la regla se define un rango para cada uno de dichos parámetros, de modo que, si allí se encuentran, inmediatamente se arrojará el resultado correspondiente. Para este caso, el SE encuentra la siguiente regla:

```
recomendar("3 y 4.5 horas",Dificultad,Perdida):-  
    Dificultad>=4,  
    Dificultad<=8,  
    Perdida>=40,  
    Perdida<=60.
```

Como se evidencia, los valores ingresados en la consulta están contenidos en el rango definido para cada variable, por ende, los valores de respuesta o recomendación coinciden con los definidos en la regla.

Consulta 2:



 sugerir(X)  

-¿Qué nivel de dificultad considera que tiene el examen?
-Indique un valor numérico entre 0 y 10 (puede ser decimal)
dificultad: 0-2 muy facil, 2-4 facil, 4-8 intermedio, 8-10 complejo

-¿Generalmente qué porcentaje de pérdida tiene asociado el examen?
-Indique un valor numérico entre 0 y 100 (puede ser decimal)
pérdida: 0-20 muy baja, 20-40 baja, 40-60 regular, 60-100 alta

Según los datos suministrados, se recomienda dedicar entre :
X = "6.5 y 9.5 horas"

Para este caso, el sistema actúa de manera similar al anterior caso. Aquí se encuentra la siguiente regla:

```
recomendar("6.5 y 9.5 horas",Dificultad,Perdida):-  
    Dificultad>=8,  
    Dificultad<=10,  
    Perdida>=60,  
    Perdida<=100.
```

Como se evidencia, los valores ingresados en la consulta están contenidos en el rango definido para cada variable, por ende, los valores de respuesta o recomendación coinciden con los definidos en la regla.