

## 9.- Verificación y validación de los modelos de simulación

### 9.1- Algunos consejos para la construcción del modelo

Al modelar un sistema, hay que tener en cuenta que:

- El modelo de simulación deberá hacerse tan sencillo como sea posible (pero no más). De este modo es mayor la comprensión del modelo y de sus resultados, y disminuye el coste de desarrollo, depurado y mantenimiento del código de programación, así como el tiempo requerido para su ejecución.
- Si el modelo de simulación se divide en submodelos, deben modelarse todos con un nivel similar de complejidad.
- El modelo se desarrolla siempre con un propósito determinado, y deben revisarse sus hipótesis si se desea emplear para otro distinto. El modelo debe estar documentado con las hipótesis acerca del sistema en que se basa.
- Los modelos computacionales suelen ser más precisos en la comparación de sistemas alternativos que en la obtención de respuestas absolutas.
- Al realizar el modelo de simulación conviene disponer de la máxima información posible acerca del sistema. Puede ser útil conversar con los "expertos" que conocen el sistema, examinar las teorías existentes acerca del mismo a fin de no "reinventar la rueda", realizar observaciones experimentales, etc.
- Durante la construcción del modelo, puede invitarse a especialistas en el sistema a que examinen datos de salida, procedentes unos del modelo y otros del sistema real, y que intenten, sin conocer la respuesta, identificar (razonadamente) cuales provienen del modelo y cuales del sistema. Esta información puede emplearse para mejorar el modelo.

### 9.2- Verificación

La *verificación* consiste en la comprobación de que el programa de simulación esta libre de errores y se comporta de modo consistente con el modelo. La *validación* consiste en comprobar que el modelo supone una aproximación suficientemente próxima de la realidad en la aplicación en concreto. Aunque los conceptos de verificación y validación son conceptualmente distintos, en la práctica, cuando la salida del programa es incomprensible no siempre esta claro si es debido a un fallo en la programación, a un fallo conceptual de modelado o a que los datos de entrada son erróneos.

Para verificar el programa pueden usarse los siguientes métodos:

- Verificación manual de la lógica: ejecutar el modelo durante un periodo corto de tiempo y comprobar manualmente los resultados obtenidos.
- Comprobación módulo a módulo: verificar individualmente cada rutina, comprobando que produce los resultados esperados para todos los posibles tipos de entradas.
- Comparar con soluciones conocidas: ajustar el modelo de modo que represente un sistema de solución conocida y comparar ésta con los resultados del modelo.
- Test de sensibilidad: variar un parámetro del modelo manteniendo el resto fijos, comprobando si el modelo es sensible a esta variación.
- Excitar el modelo: dar valores extraños a los parámetros del modelo y comprobar si el modelo se comporta de manera extraña. Algunos errores difíciles de localizar se desenmascaran excitando de este modo el modelo.

### 9.3- Validación

Una de las fases más importantes en la realización de la simulación de un sistema real, es la *validación* del modelo: determinar si éste representa de forma suficientemente precisa los aspectos relevantes de interés del sistema real bajo estudio.

El método científico consta fundamentalmente de los siguientes pasos:

- 1.- Recogida de observaciones y definición del problema.
- 2.- Formulación de una hipótesis, empleando para ello la imaginación.
- 3.- Contraste de la hipótesis mediante experimentación. Si es preciso, reformular la hipótesis.
- 4.- Confirmar que la hipótesis es igualmente válida para cualquier persona razonable.
- 5.- Definir métodos capaces de refutar la hipótesis mediante otras.

Las teorías científicas deben superar todas las pruebas anteriores. Se dice entonces que, por el momento, la hipótesis no ha podido ser refutada.

Una hipótesis o teoría nunca puede ser *probada* (K. R. Popper, 1969), sólo podemos *refutarla* o *fallar en el intento de refutarla*. Existe siempre la posibilidad de que una teoría existente sea reemplazada en el futuro por otra mejor (por ejemplo, la teoría de Newton fue reemplazada por la de Einstein). Las teorías y modelos pueden cambiar (la “verdad” tiene naturaleza relativa, no absoluta), pero los hechos permanecen inalterables.

Aunque una teoría nunca pueda ser probada, puede juzgarse su robustez a partir de su capacidad para sobrevivir a experimentos rigurosos y persistentes dirigidos a refutarla.

Al contrario que los investigadores fundamentales, la mayoría de los ingenieros dedicados al modelado pretenden realizar y validar modelos prácticos en un tiempo razonable y con un costo razonable, no pretendiendo llegar “al fondo de las cosas” a cualquier precio.

Las teorías científicas son formuladas de tal manera que pueden ser testadas en condiciones ideales en un laboratorio, por el contrario, normalmente no pueden realizarse experimentos en condiciones ideales de los sistemas reales a modelar. En general, los modelos no son lo suficientemente robustos como para resistir los rigurosos test de validación que se aplican a las teorías científicas.

Por lo general, existe una relación entre el tipo de sistema a modelar, la finalidad del modelado y el tipo de modelo matemático.

Por ejemplo, el comportamiento de los circuitos eléctricos está muy estudiado y bien entendido, con lo cual, pueden realizarse modelos para diseñar un sistema real. Si el modelo funciona correctamente, es muy probable que el modelo real también funcione correctamente. Esto es también cierto para algunos sistemas mecánicos. Este tipo de modelos matemáticos contiene normalmente ecuaciones algebraicas y ecuaciones diferenciales conteniendo sólo derivadas respecto del tiempo.

Sin embargo, esto no es cierto para sistemas químicos, ya que son muchos los factores relevantes que influyen simultáneamente sobre una reacción química, con lo cual no puede realizarse un modelo que sea válido para un conjunto amplio de experimentos. Con el nivel de conocimiento actual, no es posible realizar predicciones fiables en la fase de diseño: el comportamiento de un modelo teórico de un proceso químico puede ser significativamente diferente al comportamiento real que se observe posteriormente en el sistema real.

La finalidad de los modelos de este tipo de sistemas es el control más que el diseño. Normalmente mediante razonamientos teóricos se establece la estructura de las ecuaciones, obteniéndose los parámetros del modelo mediante ajuste a partir de los datos experimentales. A partir de este modelo puede diseñarse, normalmente de forma satisfactoria, el sistema de control de la planta, ya que los controladores tienen la tendencia de reducir la sensibilidad del sistema a las variaciones en los parámetros. Este tipo de modelos matemáticos contienen normalmente ecuaciones algebraicas y ecuaciones diferenciales conteniendo derivadas respecto del tiempo y de las coordenadas espaciales.

Cuando se pretenden modelar sistemas biológicos, la complejidad es tal que los modelos en base al conocimiento teórico actual no contienen la suficiente validez interna como para poder predecir comportamiento. El modelo puede ser usado para analizar el comportamiento del sistema: establecer relaciones causa-efecto entre las variables contempladas por el modelo. Normalmente los modelos suelen constar ecuaciones algebraicas y diferenciales respecto del tiempo.

En el caso de los sistemas económicos, el conocimiento actual ni siquiera permite construir modelos para analizar las relaciones causa-efecto. Los modelos se construyen de manera inductiva, de modo que la coincidencia entre los resultados del modelo y los experimentales no indica que la estructura interna del modelo represente la estructura interna del sistema real. Los modelos permiten, al menos, predecir en cierta medida la evolución futura del sistema real y con esa finalidad se construyen. El modelo suele estar compuesto de ecuaciones en diferencias.

Finalmente, existen también sistemas en los que ni siquiera esto es posible, en los cuales tan sólo se puede especular acerca de los posibles futuros y, tal vez, asociar una cierta probabilidad a cada uno de ellos. Entre estos están los sistemas sociales y psicológicos, en los cuales se trata de modelar el comportamiento humano. En ocasiones la propia comunidad, conocedora de estas predicciones, modifica su comportamiento de acuerdo con esta información, lo cual puede hacer que finalmente no se satisfagan. Por ejemplo, si el modelo predice un desastre, la comunidad pondrá los medios para evitarla, pudiendo quizá finalmente evitarse. Los modelos suelen estar compuestos de ecuaciones algebraicas.

En el contexto del modelado de los sistemas de colas, gestión de recursos, etc. ... que nos ocupa, las aproximaciones más comúnmente realizadas al modelar son:

- Al definir el contorno del sistema a modelar, se ignoran todos los factores externos al sistema excepto los que se consideran entradas al sistema.

- Con frecuencia se aproximan funciones altamente no lineales por otras más sencillas. Esta aproximación se hace de modo que sea “válida” en la región donde se espera que opere el modelo. Si se fuerza al modelo a operar fuera de esta región, donde la aproximación es aceptable, el programa debería imprimir un mensaje de aviso.
- Las distribuciones de probabilidad que aparecen en el sistema real son conocidas sólo de manera aproximada, con lo cual, frecuentemente, se aproximan por distribuciones teóricas sencillas. Un caso extremo es reemplazar una variable aleatoria por una constante.
- Frecuentemente se simplifica el modelo suponiendo la independencia estadística entre variable aleatorias.
- La aproximación más común es la agregación, que consiste en considerar un conjunto de entes como una unidad. Ejemplos típicos de agregación son:
  - Agregación temporal: se trata a un intervalo de tiempo como una unidad, por ejemplo, se considera que todos los eventos ocurridos en un día han sucedido simultáneamente en el mismo instante.
  - Agregación de recursos: se tratan varios recursos con un unidad. Por ejemplo, un sistema de computación con dos cpu en paralelo puede tratarse como un sistema con una velocidad el doble de rápida que una sola cpu. Un conjunto de máquinas en serie por las que tiene que pasar un producto puede considerarse como una única máquina con capacidad igual a la menor de las de las máquinas en serie.
  - Agregación de edades: en vez de considerar todas las posibles características, considerarlas agrupadas.
- Considerar que los parámetros del modelo no varían con el tiempo. Esta aproximación es razonable si la velocidad de cambio es muy pequeña comparada con el periodo de tiempo de interés.

Si es posible, conviene comprobar empíricamente las hipótesis realizadas en las fases iniciales del desarrollo del modelo, por ejemplo, las referentes a las distribuciones de entrada, usando el test chi-cuadrado.

El *análisis de la sensibilidad* del modelo a pequeños cambios en sus parámetros consiste en determinar en qué medida varía la salida del modelo al producirse un pequeño cambio en uno de los parámetros de entrada. Si la salida es particularmente sensible a algún parámetro, deberá redoblar el esfuerzo en su estimación.

Posiblemente el test más definitivo en la validación de un modelo es establecer con qué precisión se ajustan los resultados de la simulación al comportamiento real del sistema (o de alguno similar existente). Para ello, no parece demasiado lógico realizar un test de la hipótesis de que el sistema y el modelo son "lo mismo", ya que, al ser el modelo una aproximación del sistema, válida para los fines del estudio, la hipótesis será a priori falsa. Es más práctico preguntarse en qué medida las diferencias entre el modelo y el sistema son lo suficientemente significativas como para afectar a las conclusiones derivadas del modelo.

Se han desarrollado algunos métodos estadísticos para comparar los datos de salida del modelo de simulación con los correspondientes del sistema real, sin embargo, la comparación no es sencilla, ya que los procesos de salida suelen ser no estacionarios (las distribuciones de las sucesivas observaciones cambian con el tiempo) y autocorrelacionados (las observaciones en el proceso están correlacionadas unas con otras). Así pues, los métodos estadísticos clásicos basados en observaciones IID, no son directamente aplicables.

Supongamos que deseamos validar un modelo de simulación de un banco, relativo al tiempo medio de espera por cliente, durante su hora de máxima actividad: entre las 12 de la mañana y la 1 de la tarde. Si se dispone de datos reales acerca del número de clientes en el banco a las 12 de la mañana, los intervalos entre llegadas sucesivas de clientes, los tiempos de servicio y el tiempo de espera en la cola de cada cliente que llega y completa su espera entre las 12 y la 1, entonces, mejor que ejecutar la simulación generando los intervalos entre llegadas y los tiempos de servicio, a partir de las distribuciones ajustadas de los datos experimentales, es preferible suministrarle al modelo los tiempos entre llegadas y de servicio reales observados e inicializarlo con el número real de clientes en el banco a las 12. Comparando el modelo y el sistema de esta manera, obtenemos una mejor estimación de la diferencia entre ambos.

Debe usarse un conjunto de observaciones para identificar y ajustar el modelo y otro conjunto independiente de observaciones para validar el modelo. Si se usa un mismo conjunto de observaciones para construir el modelo y para validarlo, modificando su estructura o parámetros si la coincidencia no es buena y volviendo a validar, cabría preguntarse si se está construyendo un modelo válido para el sistema en general o, por el contrario, un modelo que es sólo representativo de este particular conjunto de datos.

Supongamos que, en una simulación con terminación, recogemos  $m$  conjuntos independientes de datos del sistema y  $n$  conjuntos independientes de datos de salida del modelo. Sea  $X_j$  el promedio de las observaciones del  $j$ -ésimo conjunto de datos del sistema e  $Y_j$  el promedio de las observaciones del conjunto  $j$ -ésimo de datos del modelo. Los  $X_j$ 's son variables aleatorias IID con media  $\mu_X = E(X_j)$  y los  $Y_j$ 's son variables aleatorias IID (suponiendo que los  $n$  conjuntos de datos del modelo son producidos mediante  $n$  replicaciones independientes) de media  $\mu_Y = E(Y_j)$ .

Puede compararse el modelo y el sistema construyendo un intervalo de confianza para  $\xi = \mu_X - \mu_Y$  empleando el método de los emparejamientos descrito en la sección anterior. Supongamos que hemos construido un intervalo con un  $100(1 - \alpha)\%$  de confianza para  $\xi$  y que éste no contiene al cero, entonces se dice que la diferencia entre  $\mu_X$  y  $\mu_Y$  es *estadísticamente significativa*. En cambio, si el intervalo de confianza contiene al cero se dice que la diferencia entre  $\mu_X$  y  $\mu_Y$  no es estadísticamente significativa y que las diferencias observadas entre  $\mu_X$  y  $\mu_Y$  son debidas a fluctuaciones estadísticas.

El hecho de que la diferencia entre el sistema y el modelo sea estadísticamente significativa no significa que aquel no sea una representación válida de éste. Por ejemplo, si  $\xi = 1$ , pero  $\mu_X = 1000$  y  $\mu_Y = 999$ , entonces esta diferencia probablemente no tendrá consecuencias prácticas. Se dice que la diferencia entre el modelo y el sistema es *significativa prácticamente* si la diferencia es lo suficientemente grande como para invalidar las inferencias que puedan hacerse acerca del sistema usando el modelo.