

Capítulo 1

Introducción a la Teoría de Sistemas

Capítulo 1

Introducción a la Teoría de Sistemas

Los objetivos de este capítulo son:

- Conocer todos los conceptos relacionados a la Teoría de Sistemas.
- Reconocer la importancia de la Teoría General de Sistemas dentro del desarrollo de la Ingeniería de Sistemas.
- Aplicar estos conceptos en ejemplos simples.

1.1 Introducción a los Sistemas

Uno de los conceptos más importantes de hoy en día es el de **sistemas**. Esta palabra la utilizan muchas personas y muchas veces, no estamos claros en qué es lo que significa. Una computadora, un avión, un ser humano, un automóvil, son ejemplos de sistemas, cada uno de ellos con objetivos y metas diferentes.

Los sistemas están definidos en todas partes del universo, es decir, son *ubicuos*¹. El sistema más grande es el universo y el más pequeño es el átomo. Todos los sistemas del universo tienen un nivel de organización y de jerarquía; unos se encuentran dentro de otros creando árboles de sistemas. Por ejemplo, el ser humano es un sistema y dentro de él encontramos otra serie de subsistemas: el sistema circulatorio, el sistema digestivo, el sistema respiratorio, entre otros.

Los sistemas aparecen primero en forma natural, pero con la aparición del hombre, una variedad de sistemas hechos por él comienza a existir. El ser humano vive y trabaja dentro de los sistemas. Su tecnología ha producido sistemas físicos complejos. A pesar de todos los avances, todavía la ciencia de los sistemas está en evolución. Sólo recientemente comenzamos a entender la **estructura** fundamental y las características de los sistemas naturales y de aquellos diseñados y producidos por el ser humano en un sentido científico. Puede definirse como estructura de un sistema a la relación entre sí de sus partes o entidades, donde estas relaciones pueden ser de espacio, tiempo, jerarquía y propiedades físicas / lógicas o de toma de decisiones.

Un sistema puede incluir tanto a personas como maquinarias, así como otros elementos ya sean estos físicos o abstractos. Los sistemas sociales que representan interacción entre personas incluyen la familia, empresas de negocios, economía nacional y las relaciones internacionales. Algunos de nuestros más importantes sistemas incluyen la interacción entre los elementos físicos y sociales. La tecnología moderna está rodeada por sistemas físicos, socio-económicos, organizacionales, de comunicación y muchos

¹ Están en todas partes.

otros.

Las leyes que gobiernan la forma en que los sistemas cambian, crecen y fluctúan tratan de explicar a través de la matemática de ecuaciones diferenciales y transformaciones de Laplace. Tal matemática es difícil, ya que es accesible solamente a expertos matemáticos. Aun aplicando todo el esquema del conocimiento matemático a la teoría de sistemas, la Matemática resulta débil al ser confrontada con preguntas importantes sobre los sistemas, ya que ésta es inadecuada para tratar con la realidad de muchos de los problemas del mundo. Por tales motivos, surge un nuevo enfoque matemático *discreto* que permite entender más fácilmente los conceptos básicos de los sistemas. La tarea principal de este libro es presentar este nuevo enfoque discreto y dinámico para analizar y diseñar sistemas.

Otros ejemplos de sistemas son:

- Un reloj que marca la hora.
- Una podadora es un conjunto de partes que forman un sistema para cortar el césped.
- Un gobernador de gasolina y un motor al que éste es acoplado forman un sistema para proporcionar potencia a una velocidad constante que debe mantener el motor.

1.2 Definición de Sistemas

El término "sistema" proviene de la palabra griega "systema", que significa *un todo organizado*. Algunas de las variadas formas sencillas de definir un sistema son:

1. Combinación de elementos o partes que forman un ente complejo o una unidad buscando un objetivo.
2. Conjunto de componentes y eventos relacionados que interactúan unos con otros para ejecutar una tarea.

Se puede apreciar que en estas definiciones los aspectos primordiales son conjuntos de elementos, relaciones, comunicación y objetivo común. Luego de este breve estudio sobre el significado de la palabra **sistemas**, aceptaremos como la definición más completa la siguiente:

"Una colección de entidades¹ (hombres, máquinas, facilidades, energía, materiales, órdenes, dinero, información, etc.) debidamente estructuradas, integradas y coordinadas para alcanzar una meta común y unidas a través de una red de comunicación".

A continuación, se mencionan algunos ejemplos:

- Sistemas fluviales (ríos con sus afluentes).
- Sistemas de transporte (urbano y de ferrocarril).
- Principios o doctrinas en un campo particular del pensamiento.
- Sistema capitalista o socialista.
- Una compañía de negocios.
- Sistemas de producción.
- Sistemas de mercadeo.
- Sistemas de numeración o medición.



Sistema de Transporte

1.3 Componentes de un Sistema

Los sistemas están compuestos de los siguientes componentes o elementos importantes:

Entidades:	Son los componentes o partes operativas únicas del sistema. Cada componente del sistema puede asumir una variedad de valores para describir el estado del sistema como conjunto por la acción de control y una o más restricciones. También se les conoce como <i>objetos</i> .
Atributos:	Son las propiedades, características o manifestaciones apreciables de las entidades del sistema. Un atributo puede pertenecer a una o más entidades.
Relaciones:	Son los enlaces o vínculos entre entidades del sistema. En otras palabras, son asociaciones de dos o más entidades.
Actividades	Acciones que ejecuta internamente el sistema. También se les conoce como actividades endógenas o funciones.
Frontera:	Límite del sistema.
Medio Ambiente:	Todo lo externo al sistema. Las actividades que se llevan a cabo en el medio ambiente se le conocen como actividades exógenas. En el medio ambiente pueden existir n número de sistemas que se pueden comunicar con el sistema en cuestión.
Estado:	Es la descripción de las condiciones que prevalecen en <i>un momento dado</i> en los diferentes componentes de un sistema.
Unidad de Tiempo:	Se utiliza como medida del funcionamiento del sistema, se aplica globalmente y a través de toda la ejecución del sistema.
Cambios:	Los cambios son los pasos de un estado hacia otro estado y pueden ser: Reactivos: Una reacción es un evento causado determinísticamente por otro evento. De respuesta: Es cuando un evento puede crear una causa, la cual no es suficiente para provocar un efecto. Autónomos: Son eventos determinados por sí mismos, los cuales no necesitan un evento predecesor.

En la figura 1.1 se presentan los elementos de un sistema. Existen otros elementos que se deben considerar, pero se presentaron sólo los más importantes.

1.4 Propiedades de los Sistemas

Los sistemas poseen una serie de propiedades que rigen su estructura y comportamiento. A continuación, se presentan cuatro de las más importantes:

1. Las propiedades y comportamientos de cada componente del conjunto tienen un efecto vital sobre las propiedades y comportamiento del conjunto como un todo.
2. Las propiedades y comportamientos de cada componente dependen de las propiedades y comportamientos de por lo menos un componente del conjunto.

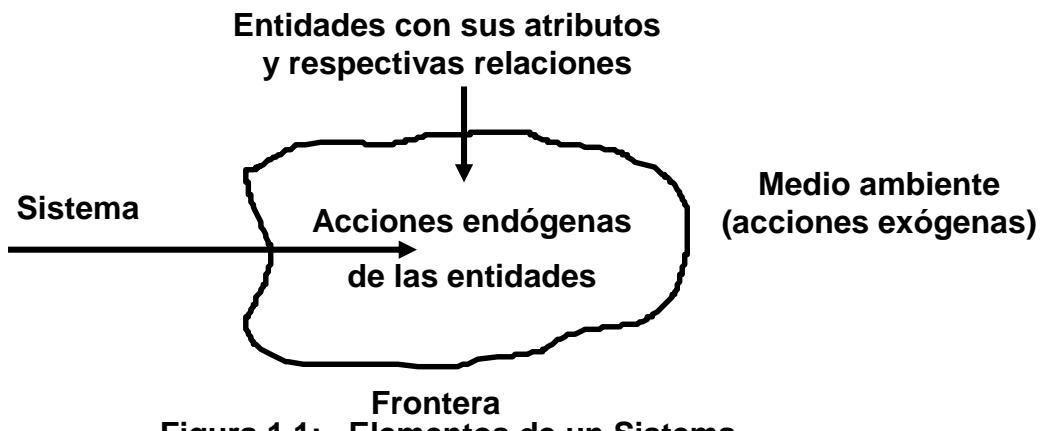


Figura 1.1: Elementos de un Sistema

3. Cada posible subconjunto de componentes tiene las dos propiedades anteriores; los componentes no pueden dividirse en subconjuntos independientes.
4. Las propiedades dadas aseguran que un conjunto de componentes, comprenden un sistema siempre que tengan algunas características o comportamiento patrón que no puede ser mostrado por cualquier subconjunto.

1.5 Algunas Clasificaciones de los Sistemas

Los sistemas se pueden clasificar tomando en cuenta:

- **Tamaño:** Poseen una cantidad pequeña, considerable o excesiva de componentes.
- **Complejidad:** Diversidad que va desde fácil de entender hasta los de un nivel muy variado.
- **Nivel de automatización:** Manual, semiautomático y automático.

- **Nivel de Entendimiento²:**
 - **Determinístico:** Comprendemos totalmente el comportamiento del sistema.
 - **Estocástico:** No determinístico, no comprendemos perfectamente el sistema, es decir, que el comportamiento del sistema puede variar con base a ciertos rangos de probabilidad.
- **Discretos:** Un sistema discreto abarca una o más variables que son conocidas sólo en instantes separados o puntuales en el tiempo.
- **Continuos:** El valor del sistema, está representado mediante variables las cuales son función de un tiempo continuo y que determinan los cambios en su comportamiento.

A continuación, se presentan algunas de las clasificaciones más comunes:

1.5.1 Sistemas Naturales versus Sistemas Hechos por el Hombre

Los sistemas naturales son aquellos cuya existencia se da a través de procesos naturales y su descripción es la tarea del astrónomo, físico, biólogo, psicólogo o sociólogo, entre otros. Los sistemas naturales son muy abundantes en la naturaleza y cada organismo viviente es un sistema natural único y propio. **Ejemplos:** Organismo humano, sistema solar, flora, fauna, cuenca hidrográfica, sistema fluvial, fotosíntesis, entre otros.



Los sistemas hechos por el hombre son aquellos que han sido desarrollados por seres humanos, resultado de una gran variedad de habilidades; se formaron cuando los hombres se reunieron por primera vez en grupos para vivir y cazar juntos. **Ejemplos:** Caminos, carretas, hospitales, fábricas sistemas de contabilidad, automóvil, computadora y muchos otros sistemas.



Sistemas hechos por el hombre

1.5.2 Sistemas Estáticos versus Sistemas Dinámicos



Un sistema estático es aquel que posee una estructura sin movilidad (casi sin) o actividad interna. **Ejemplos:** Una estatua, una silla, una pared, una mesa, etc.

Un sistema dinámico combina componentes estructurales con actividad.

Ejemplos: Automóvil, trasbordador espacial, avión, una ciudad con todos sus elementos, sistema de drenaje, sistema circulatorio, el planeta tierra.

Otra forma más avanzada de ver esta clasificación es desde el punto de vista de

² Puede existir una categorización intermedia.

aprendizaje del sistema. Esto quiere decir que un sistema estático no aprende de sus acciones pasadas y un sistema dinámico aprende de sus acciones pasadas (realimentación) y reacciona con base a lo que aprende³. La diferencia básica entre un sistema estático y un sistema dinámico es que el segundo utiliza las actividades pasadas para planificar o mejorar las actividades del futuro, es decir, que se realimenta. La filosofía de los sistemas dinámicos cree que el comportamiento a través del tiempo de una organización es causado principalmente por la estructura de dicha organización (aspectos físicos y aspectos intangibles). La tendencia humana es analizar y diseñar los sistemas en forma estática y debido a ello los sistemas tienden a no ajustarse a los cambios a través del tiempo y tienden a desaparecer, ya que no han aprendido del pasado, y es por ello que el análisis debe ser enfocado en función de sistemas dinámicos.

Esta clasificación es el núcleo de este nuevo enfoque y el motivo principal de este libro. En las secciones 1.7.2 y 1.7.3 se presenta este enfoque con un poco más de detalle.

1.5.3 Sistemas **Adaptables** versus Sistemas **No Adaptables**

Un sistema adaptable reacciona al medio ambiente, en forma tal, que pueda auto ajustarse o auto modificarse de acuerdo con las reacciones impredecibles del medio ambiente, para poder cumplir con su(s) metas(s).

Además de ajustarse a los cambios ambientales, también lo hace ante errores moderados del proyecto de ingeniería o incertidumbres y compensa la eventual falla de componentes menores del sistema, aumentando, por tanto, la confiabilidad de todo el sistema. **Ejemplos:** El hombre, las plantas, los animales, el camaleón (todos hasta cierto punto, esto quiere decir, dentro de un rango).



En un sistema no adaptable un cambio significativo en el ambiente no permite que el sistema se ajuste a él y lo puede llevar a que, a corto o largo plazo, no sobreviva. **Ejemplos:** Peces fuera del agua, edificio no diseñado a prueba de terremotos, un reloj que ha sido golpeado severamente y deja de funcionar.

1.5.4 Sistemas **Manuales** versus Sistemas **Automáticos**

En un sistema manual, las acciones requieren de la intervención humana.

Ejemplos: Orfebrería, agricultura primitiva, organizaciones de negocios, partidos políticos, etc.

³ Esto crea una diferencia entre *dinámica simple* y *dinámica compleja*. Ver sección 1.7

En realidad, no existen sistemas automáticos porque la mayoría de los sistemas son semiautomáticos, donde la máquina lleva a cabo acciones que son controladas por el hombre. Indiscutiblemente que, a través del tiempo, las máquinas ejecutarán más tareas que el hombre.

Ejemplos: El operador y la consola de la computadora, brazo mecánico, microfilmación, ensamblaje de automóviles, el perro Aibo de la Sony que reconoce al amo y aprende de la interacción con su dueño, entre otros.



1.5.5 Sistemas Abiertos versus Sistemas Cerrados

Un sistema abierto es aquel que se comunica con su medio ambiente o entorno a través de su frontera. Esto quiere decir que las actividades exógenas (externas al sistema) actúan sobre las actividades endógenas (internas) del sistema y viceversa, lo que significa que el sistema cuenta con un mecanismo de acoplamiento con su entorno.

Ejemplos: Un ciudadano dentro de la sociedad, alimento perdurable sin envoltura, termostato del acondicionador de aire.

Un sistema cerrado no se comunica con su medio ambiente, o sea que la frontera del sistema impide que las actividades endógenas interactúen con las actividades exógenas.

Ejemplos: Un reloj de pulsera, alimento perdurable debidamente sellado.

1.5.6 Sistemas Físicos y Conceptuales

Los sistemas físicos son aquellos gobernados por componentes reales que ocupan un espacio. Por otra parte, los sistemas conceptuales conforman una organización de ideas, un conjunto de especificaciones y planes, una serie de conceptos abstractos.

Ejemplos: Una persona constituye un sistema físico, las leyes que rigen un país, un sistema conceptual, o sea, la forma como se guardan los datos dentro de un disco magnético de una computadora.

1.6 Sistemas de Ciclo Abierto y de Ciclo Cerrado⁴

Los sistemas también pueden clasificarse como sistemas de "**ciclo abierto**" o sistemas de "**ciclo cerrado**" (también conocido como de **realimentación o sistema dinámico**)⁵.

Un sistema de ciclo abierto está caracterizado por salidas que responden a entradas, pero donde las salidas están aisladas y no tienen influencia en las entradas. Por lo tanto,

⁴ Nótese que existe una diferencia entre *sistemas abiertos* y *sistemas de ciclo abierto*.

⁵ Debido a la importancia de esta clasificación, se ha colocado como un punto aparte en este capítulo.

el sistema no está al tanto de su rendimiento. En un sistema de ciclo abierto, la acción pasada no controla la acción futura. Un sistema de ciclo abierto no observa y reacciona a su rendimiento. Un automóvil es un sistema de ciclo abierto ya que no se controla a sí mismo, no sabe dónde estuvo en el pasado ni a dónde irá en el futuro. Un reloj, controlado por sí mismo, no advierte su propia inexactitud ni se ajusta por sí sólo, por eso es un sistema de ciclo abierto.



Figura 1.2: Sistema de Ciclo Abierto

Otro ejemplo práctico lo constituye una lavadora de ropa doméstica. El remojo, lavado y enjuague en la lavadora se cumplen por tiempos. La máquina no mide la señal de salida, es decir, la limpieza de la ropa. El control de tráfico con señales accionadas en función de tiempos, es otro caso de control de ciclo abierto.

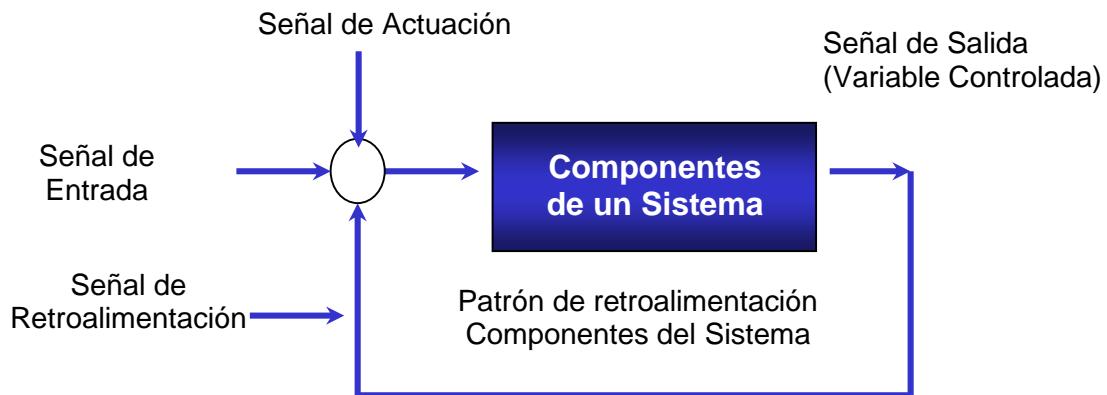


Figura 1.3: Sistema de Ciclo Cerrado

Un sistema de ciclo cerrado, el cual muchas veces es llamado sistema de realimentación⁶, es influenciado por su comportamiento de ciclo cerrado. Un sistema de realimentación trae resultados de una acción pasada que afectarán a las acciones futuras del sistema (véase la figura 1.3).

Por ejemplo, un control de temperatura del ambiente de una habitación: midiendo la temperatura efectiva de la habitación y comparándola con la temperatura de referencia (temperatura deseada), el termostato conecta o desconecta los equipos de calefacción o refrigeración, de modo que la habitación se mantiene a una temperatura confortable,

⁶ O retroalimentación.

independientemente de las condiciones del exterior e interior.

En el cuerpo humano también observamos que la temperatura y la presión se mantienen en valores constantes por medio de una realimentación fisiológica. De hecho, la realimentación cumple una función vital: hace al cuerpo humano relativamente insensible a perturbaciones externas, permitiéndole desenvolverse adecuadamente en un medio ambiente cambiante.

Un sistema de realimentación controla la acción basado en los resultados de acciones previas. Existen dos tipos de realimentación: la positiva y la negativa. La realimentación positiva refuerza la dirección en la cual el sistema se está moviendo.

La realimentación que pretende el ajuste, amortiguamiento y reducir las fluctuaciones alrededor del estándar deseado del comportamiento del sistema una vez que se detecta la no-consecución de las metas del sistema, se denomina realimentación negativa.

1.6.1 Ejemplo de un Sistema de Realimentación Negativa

El sistema de calefacción de una casa se controla por un termostato el cual responde al calor previamente producido por el horno. El calor producido por el sistema busca como meta llegar a la temperatura deseada.

Un reloj y su dueño forman un sistema de realimentación negativo cuando el reloj se compara con el tiempo correcto como una meta, y se ajusta para eliminar errores.

1.6.2 Ejemplo de un Sistema de Realimentación Positiva

Las bacterias se multiplican para producir más bacterias las cuales incrementan la tasa en que las nuevas bacterias son generadas. En este sistema de realimentación positivo, la tasa de generación de nuevas bacterias depende de las bacterias acumuladas con el crecimiento de bacterias.

Este tema es tratado con mayor amplitud en los capítulos 2 y 5 de este libro, ya que es el tema central de la teoría de sistemas dinámicos.

1.7 Clasificación de los Sistemas con base a su Complejidad

Kenneth Boulding propuso la siguiente clasificación con base a la complejidad que exhiben los sistemas. Esta clasificación es muy importante ya que nos permite entender la jerarquía de los sistemas. El orden de los diez (10) niveles es de menor a mayor complejidad⁷:

1.7.1 Estructura Estática

Este nivel forma la geografía básica del universo. Posee una serie de sistemas que se han mantenido iguales a través de millones de años. A la fecha este nivel todavía no se

⁷ El noveno nivel lo introduce el autor.

comprende totalmente y los científicos tratan de presentar una explicación a todos sus fenómenos.

1.7.2 Sistemas Dinámicos Simples

Este nivel encierra una serie de disciplinas desde el enfoque de sistemas dinámicos, tal como lo definimos en la sección 1.5.2, hasta el ángulo de movilidad, acción.

1.7.3 Sistemas Dinámicos Complejos⁸

Este nivel recoge todos los sistemas desde el otro ángulo, definido en la sección 1.5.2, desde el ángulo de realimentación o sistemas de ciclo cerrado. El sistema aprende de sus acciones pasadas, es decir, que posee un mecanismo de control.

1.7.4 Célula

Este nivel es el primero que caracteriza la *vida* como ingrediente trascendental en la complejidad de los sistemas.



1.7.5 Flora

Este nivel recoge todo lo concerniente a la botánica del universo. Las células forman sociedades de células con componentes diferenciados, pero mutuamente dependientes.



1.7.6 Fauna

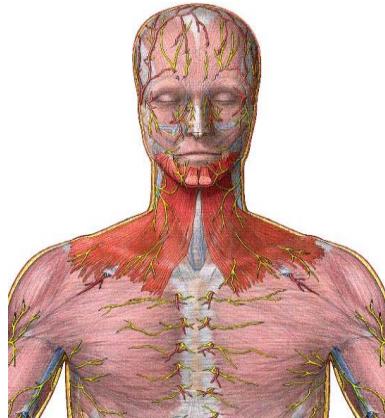
Aquí se encuentran todos los animales del planeta. Se desarrollan dispositivos de entrada, entrada-salida y salida como los sentidos. También se desarrollan sistemas nerviosos conectados a cerebros que organizan información.



1.7.7 Ser Humano

⁸ En la sección 1.5.2 se menciona que los sistemas dinámicos poseen dos niveles de complejidad.

Este nivel recoge el nivel de inteligencia humana. En este nivel también incluimos toda la fauna y la flora; esto quiere decir que este nivel representa vida en todos los niveles.



1.7.8 Sociedad

Este nivel tiene que ver con los seres humanos como grupo. Recoge todo lo concerniente a la teoría de comunicación, comportamiento grupal, principios, ética, cultura, emociones y el conocimiento humano.



1.7.9 Inteligencia Artificial

Este nivel recoge todos los sistemas hechos por el hombre que poseen una semejanza con la inteligencia humana. Algunas de las disciplinas dentro de la inteligencia artificial son lenguaje natural, reconocimiento de imágenes, sistemas basados en el conocimiento, robótica, sistemas expertos etc.



1.7.10 Desconocido

Este nivel recoge todo lo que a la fecha se desconoce. Los niveles, tal como podemos apreciar, se rigen por una serie de principios físicos, químicos y biológicos. A partir del nivel célula, comienza a presentarse el concepto de vida y a partir del nivel de planta comienzan a surgir diferentes niveles de inteligencia.

1.8 Ingeniería de Sistemas

La Ingeniería de Sistemas analiza, diseña, construye, examina, opera, controla y evalúa sistemas que se usan para representar y resolver problemas y procesos en áreas industriales, de desarrollo, de administración, relaciones humanas entre otros con el propósito de que sean más eficientes⁹. Como disciplina, recoge una serie de principios

⁹ La Ingeniería de Sistemas cuenta hoy con el “Systems Engineering Body of Knowledge”.

científicos y de tecnologías que puedan mejorar el desempeño de los sistemas. Podríamos decir que básicamente es una ingeniería de enlace entre las demás disciplinas para permitir el adecuado enfoque de los sistemas. Se basa en un conocimiento formal de las interacciones entre las partes de un sistema. Ella considera el contenido de nuevos conocimientos, luego planea y participa en la acción de los nuevos proyectos y en el programa completo de los proyectos encaminados a las aplicaciones. Tiene en cuenta las necesidades de los consumidores y determina la forma de satisfacerlas, con la ayuda de todos los conocimientos, ya sean anteriores o recientes. Por lo tanto, la ingeniería de sistemas actúa en el espacio comprendido entre la investigación y las operaciones mercantiles. Ingeniería porque su énfasis está en las operaciones de conceptos cuantitativos o problemas concretos, mientras que la palabra "sistema" describe su tendencia a analizar problemas desde su punto global.

La Ingeniería de Sistemas es un arte porque a veces no existen procedimientos sistemáticos que determinen exactamente donde trazarse la frontera entre el sistema y su ambiente, cuáles tienen que ser los componentes del sistema, o cuál debe ser el objetivo del sistema. Es una ciencia porque posee sus propios axiomas y métodos cuantitativos, investigación de operaciones y usa técnicas y conceptos matemáticos como la estadística, probabilidad, etc.

La Ingeniería de Sistemas es una especialidad y una generalidad, pues cada vez que se resuelve un problema, se aprende algo de una nueva disciplina o profesión. Ha tenido un éxito comprobado en una gran variedad de disciplinas, sobre todo, en el análisis de sistemas a gran escala, tales como transporte, planeación urbana, administración, contabilidad, educación, computación, ingeniería espacial, etc.

La metodología de sistemas se puede conceptualizar utilizando una serie de etapas conocidas como el ciclo básico de un sistema. Las etapas representan la evolución del sistema desde su planeación inicial, hasta su implementación.

Las intercomunicaciones, la compatibilidad, el efecto del uno sobre el otro, los objetivos del todo, la relación del sistema con los usuarios, y la factibilidad económica debe recibir aún más atención que las partes, así el resultado final podrá ser exitoso.

Ejemplo: Un sistema telefónico no es solamente cables, amplificadores, postes de luz y aparato telefónico considerados individualmente. Estas 4 partes se encuentran interconectadas y coordinadas con el fin de brindar un servicio de comunicación al público usuario.

Algunos de los beneficios de esta ingeniería son:

- La tendencia a cuantificar. Una de las contribuciones más importantes de la Ingeniería en Sistemas de hoy en día es la tendencia a cuantificar el valor de las alternativas, componentes o soluciones del proyecto. El uso ya mencionado de modelos matemáticos en éste, obliga al analista a comparar alternativas bajo un criterio o medida de valor común.
- La tendencia a resolver problemas a gran escala. Se da resolución a problemas

que constan de muchos componentes, frecuentemente difíciles de identificar y altamente interconectados.

- La oportunidad de analizar y diseñar tomando en cuenta todos los componentes existentes y entendiendo claramente sus interrelaciones, actividades, etc.

Cuando ha sido necesario obtener resultados numéricos en estos estudios, la computadora se convirtió en el arma principal del ingeniero en sistemas. Esto no significa que se debe utilizar la computadora en todas las aplicaciones de ingeniería de sistemas. En todos aquellos casos en los cuales el uso de la misma se haga necesario, desde un punto de vista técnico, se deberá estudiar el problema más a fondo para determinar, si el valor de los resultados numéricos justifica el costo de utilizar la computadora.

1.9 Objetivos de la Ingeniería de Sistemas

A continuación, se presentan algunos de los objetivos más importantes:

- Estudiar y comprender los sistemas para anticipar o mejorar su comportamiento.
- Proporcionar a la gerencia toda la información que sea posible y necesaria para una guía y control del programa general de desarrollo.
- Formular planes de largo alcance y objetivos, como un marco para enlazar entre sí los proyectos individuales.
- Balancear el programa general de desarrollo a fin de asegurar el progreso a lo largo de todas las líneas de demandas, y haciendo al mismo tiempo el mejor empleo del desarrollo de la mano de obra y de otros recursos.
- Desarrollar los objetivos y los planes para proyectos particulares y hacerlos consistentes con los objetivos más lejanos. Conocer las necesidades actuales de la organización.
- Prever con anticipación las necesidades futuras a fin de estar completamente preparados para el momento en que se requiera una acción.
- Tener siempre presente nuevas ideas, principios, métodos y dispositivos.
- Efectuar cada una de las operaciones del proceso de la ingeniería de sistemas en la forma más eficiente posible, reconociendo que los requisitos para los detalles, la exactitud y la rapidez dependen sólo de las fases del proceso en que se esté trabajando.
- Integrar una gran variedad de disciplinas de diseño clave al esfuerzo de la ingeniería de diseño total.
- Planificar, analizar, diseñar, evaluar sistemas hechos por el hombre para las áreas físicas, industriales, organizacionales, sociales, etc.

1.10 Usos de los Sistemas Dinámicos

Un sistema dinámico es la aplicación de los principios y técnicas de los sistemas de control de realimentación a los problemas físicos, administrativos, organizacionales y socioeconómicos. Esto quiere decir que sistemas de realimentación significa lo mismo que sistemas dinámicos o que sistemas de ciclo cerrado.

Todo sistema posee unas metas y objetivos que debe cumplir. Adicionalmente, por ser dinámico, simple o complejo, ejecuta una serie de acciones y pasa de un escenario (estado) a otro dependiendo de ciertas condiciones. El módulo de realimentación y control le permite verificar sus metas y actuar con base a los resultados de la comparación entre la medición de la situación actual y la(s) meta(s) a lograr. (Véase la figura 1.4).

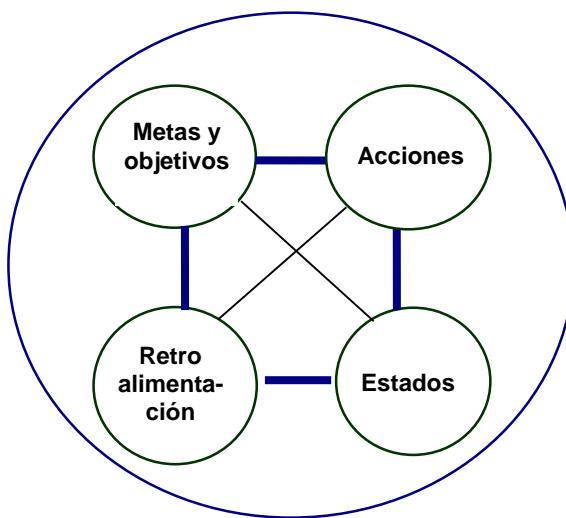


Figura 1.4: Estructura Funcional de un Sistema

Para uso administrativo, los sistemas dinámicos buscan la integración de las diferentes áreas funcionales de una organización para crear un conjunto conceptual que tenga sentido y proveer una organización al igual que bases cuantitativas para diseñar políticas organizacionales más efectivas, esto se logra por la realimentación de hechos pasados para mejorar el rendimiento futuro de la organización. Los sistemas dinámicos son el estudio de las características de información y realimentación de las actividades comerciales e industriales para mostrar como la estructura organizacional, las políticas, las acciones y decisiones, y los retrasos de tiempo interactúan para influir en el progreso de la empresa. Se trata de las interacciones de los flujos de información, dinero, órdenes, materiales, personal, equipo de una compañía comercial, de servicio, industrial o la economía nacional como un todo. Los sistemas dinámicos proveen un sencillo armazón para integrar las áreas funcionales de administración, mercadeo, producción, cuentas, investigación y desarrollo, y capital de inversión. Es un acercamiento cuantitativo y experimental para relacionar la estructura organizacional y las políticas generales para

la estabilidad y crecimiento industrial. Los sistemas dinámicos suministran una base de diseño de sistemas económicos e industriales más efectivos.

Los pasos para diseñar un sistema dinámico son:

1. Identificar el problema. Describir qué problemas tiene el sistema que requieren ser mejorados.
2. Aislar los factores que interactúan para identificar los síntomas observados.
3. Estudiar los ciclos de causa y efecto de información y realimentación que unen las decisiones a las acciones, de tal forma que la acción resultante cambie las nuevas decisiones.
4. Formular reglas, formales y aceptables, que describan como las decisiones resultan de las corrientes de información disponible.
5. Construir un modelo matemático de las reglas de decisiones, fuente de información e interacciones de los componentes del sistema.
6. Generar el comportamiento a través del tiempo del sistema descrito en el modelo (usualmente se ejecutan los cálculos por medio de computadoras).
7. Comparar todos los resultados contra todo el conocimiento pertinente disponible del sistema actual.
8. Revisar el modelo hasta que sea aceptable y sea una representación del sistema actual.
9. Experimentar con el modelo. Rediseñar dentro del modelo las relaciones organizacionales y reglas que puedan ser alteradas en el sistema actual para encontrar los cambios que mejoren el comportamiento del sistema.
10. Alterar el sistema actual en las direcciones del experimento en el modelo para mejorar el rendimiento del sistema¹⁰.

La filosofía de los sistemas dinámicos cree que el comportamiento (a través del tiempo) de una organización es causada principalmente por la estructura de la organización. La estructura no incluye solamente los aspectos físicos de un proceso de planta o producción, sino que también las normas y reglas tangibles e intangibles, que dominan la toma de decisiones de la organización. Tal estructura contiene fuentes de amplificaciones, información de realimentación similar a aquellas que se encuentran en complejos sistemas de ingeniería. Los sistemas de ingeniería y administración que contienen estas características muestran modelos de respuestas complicados para sistemas relativamente simples o de entrada variable. El análisis de un gran sistema no lineal de esta clasificación es un gran desafío aún para el ingeniero más experimentado en el sistema. El rediseño efectivo y justificable de sistema como este es aún más difícil. Las sutilezas y complejidades en el área de administración hacen estos problemas

¹⁰ El capítulo 3 presenta estos pasos con más detalle.

mucho más difíciles.

Un segundo aspecto de la filosofía de los sistemas dinámicos es el concepto de que las organizaciones son vistas más efectivamente desde el punto de vista de sus flujos, que de sus funciones separadas. Los flujos de personas, dinero, material, órdenes y capital de equipo y la integración de flujos de información pueden ser identificados en todas las organizaciones. La orientación de la estructura de los flujos hace que el analista cree las fronteras de la organización de una forma natural. Este actúa separando los componentes próximos de la organización que provocan conflictos organizacionales y la suboptimización no reconocida. Un significativo armazón de sistemas resulta al trazar cadenas de causa y efecto a través del espacio de los flujos apropiados.

1.11 Temas Relacionados con los Sistemas

1.11.1 La Cibernética

La palabra cibernética en griego se refiere a mecanismos precisos de gobierno y control. Con Platón y Ampere, es usada siempre en su sentido político-social, pero es utilizada por primera vez en referencia a la ingeniería humana por Norbert Wiener.

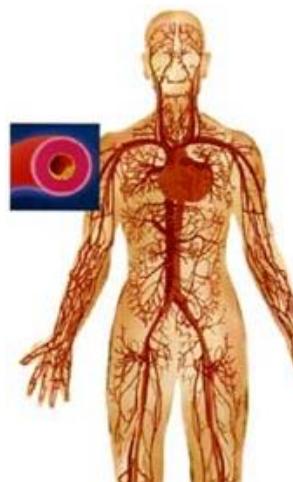
La cibernética es una disciplina íntimamente vinculada con la teoría general de sistemas, al grado en que muchos la consideran inseparable de esta, y se ocupa del estudio del mando, control, regulaciones y la gobernabilidad de los sistemas. El propósito de la cibernética es desarrollar lenguajes y técnicas que nos permitan atacar los problemas de control y comunicación en general.

La Teoría General de Sistemas y la Cibernética esencialmente estudian el mismo problema y están íntimamente ligadas, pero la distinción que podemos hacer notar es que la primera está enfocada más en la estructura y los modelos de los sistemas, mientras que la segunda está enfocada al control de las acciones de los sistemas, a cómo se comunican con otros sistemas o con sus propios elementos.

1.11.2 La Homeostasis

El concepto de homeostasis fue introducido en la fisiología en 1932 por W. Cannon, para explicar la constancia relativa de ciertas dimensiones fisiológicas. Por ejemplo, la temperatura del cuerpo de los mamíferos que se mantiene constante, frente a la temperatura cambiante del ambiente externo.

La homeostasis es la propiedad de un sistema que define su nivel de respuesta y de aceptación al contexto. Es el nivel de adaptación permanente del sistema o su tendencia a la supervivencia dinámica. Los sistemas altamente homeostáticos sufren transformaciones estructurales en igual medida que el contexto



sufre transformaciones, ambos actúan como condicionantes del nivel de evolución.

La homeostasis es posible por el uso de información proveniente del medio externo incorporada al sistema en forma de retroalimentación (feedback). El "feedback" activa el "regulador" del sistema, que, alterando la condición interna de éste, mantiene la homeostasis. Un ejemplo muy común del modo cómo funciona la homeostasis es el de un sistema de calefacción central, que mantiene a la casa en un estado estable de calor. Utiliza un termostato, que desempeña el papel de regulador y que responde al feedback referente a la temperatura del "supra sistema" exterior a la casa. Cuando la temperatura exterior desciende, el termostato actúa aumentando la temperatura dentro de la casa. La homeostasis es un mecanismo auto correctivo.

1.11.3 La Entropía

La entropía de un sistema es el desgaste que el sistema presenta por el transcurso del tiempo o por el funcionamiento del mismo. Los sistemas altamente entrópicos tienden a desaparecer por el desgaste generado por su proceso sistémico. Los mismos deben tener rigurosos sistemas de control y mecanismos de revisión, reelaboración y cambio permanente, para evitar su desaparición a través del tiempo.

La entropía, quiere decir, la máxima probabilidad de los sistemas en su progresiva desorganización, y finalmente su homogeneización con el ambiente. Los sistemas cerrados están irremediablemente condenados a la desorganización. No obstante, hay sistemas que, al menos temporalmente, revierten esta tendencia al aumentar sus estados de organización (neguentropía, información).

1.11.4 La Neguentropía

Los sistemas vivos son capaces de conservar estados de organización improbables (entropía). Este fenómeno aparentemente contradictorio se explica porque los sistemas abiertos pueden importar energía extra para mantener sus estados estables de organización e incluso desarrollar niveles más altos de improbabilidad. La neguentropía, entonces, se refiere a la energía que el sistema importa del ambiente para mantener su organización y sobrevivir. Neguentropía es entonces un sistema dinámico de retroalimentación que se trata en el capítulo 3.

La entropía tiende a desordenar el sistema, sin embargo, el sistema a través de la neguentropía puede combatir y superar esa tendencia.¹¹

1.11.5 La Sinergia

La palabra sinergia se deriva del latín *synergos*, que significa "trabajar en conjunto". Su

¹¹ Definición según la Real Academia de la Lengua Española.

aplicación en el mundo de los negocios se refiere a la habilidad de dos o más unidades o compañías para generar mayor valor trabajando en conjunto, que aquel que podrían producir trabajando en forma separada.

En pocas palabras, se logra la sinergia cuando dos o más sistemas o personas trabajan conjuntamente para crear una mejor solución de lo que ambos pudieran lograr por cuenta propia. No es tu forma o la mía, sino una forma, una más elevada.

Todo sistema es sinérgico en tanto el examen de sus partes en forma aislada no puede explicar o predecir su comportamiento global. La sinergia es, en consecuencia, un fenómeno que surge de las interacciones entre las partes o componentes de un sistema.

La sinergia no es algo que simplemente sucede. Y el fundamento para llegar hasta allá es el siguiente: Aprender a celebrar las diferencias.

Sinergia es la acción de dos o más causas cuyo efecto es superior a la suma de los efectos individuales.

1.12 Páginas Web sobre la Dinámica de Sistemas

Dinámica de Sistemas “Daedalus”

URL:<http://www.daedalus.es/AreasDSHistoria-E.php>

Descripción: El sitio nos habla sobre la dinámica de sistemas, su historia, siendo pionero el Ing. Jay W. Forrester, nos define la dinámica de sistemas como el método para la construcción de modelos de simulación para modelos complejos como los que son estudiados por las ciencias sociales, la economía o la ecología. También nos habla de sus aplicaciones, herramientas.

Dinámica de sistemas

URL: <http://www.itson.mx/dii/elagarda/apagina2001/Dinamica/dsistemas.html>

Descripción: En este sitio nos hablan de la introducción a la dinámica industrial, construcción de los modelos de la dinámica de sistemas, los elementos de la dinámica de sistemas, las estructuras elementales de los sistemas dinámicos, además de un proyecto final en el cual muestran un ejemplo de sus aplicaciones.

Revista de dinámica de sistemas

URL: http://dinamicasistemas.utalca.cl/Revista/RDS_home.htm

Descripción: La Revista de Dinámica de Sistemas es una publicación científica con arbitraje por pares. Se dirige a una audiencia amplia para comunicarle los avances en

los métodos y la perspectiva de la Dinámica de Sistemas en los ámbitos de problemas sociales, técnicos, de administración y ambientales. La revista publica avances en el modelado matemático y la simulación computacional de sistemas dinámicos de retroalimentación; avances en metodologías de análisis de políticas basados en la retroalimentación de información y la causalidad circular, etc.

Dinámica de sistemas y control

URL: <http://ib.cnea.gov.ar/~dsc/index.html>

Descripción: es un sitio que está dividido por capítulos que son introducción, variables de estado, soluciones de las variables de estado, grafos de enlaces, estabilidad de sistemas, etc.

Dynamic Systems

URL: http://en.wikipedia.org/wiki/Dynamic_system

Descripción: Descripción completa de las áreas que conforman un Sistema Dinámico

Fourth International Conference on Dynamic Systems and Applications

URL: <http://atlas-conferences.com/cgi-bin/abstract/cajw-01>

Descripción: Desarrollos en los cuales se aplicó la teoría de Sistemas Dinámicos

Seminarios sobre Control y Dinámica de Sistemas

URL: <http://www.cds.caltech.edu/>

Descripción: Página con itinerario de seminarios y artículos que están relacionados con la dinámica de sistemas.

The Math Forum Drexel

URL: http://mathforum.org/library/topics/dynamical_systems/

Descripción: Foro sobre Dynamic System

Universidad Politécnica de Catalunya

URL: <http://www.upcnet.es/~jmg2/sis3.htm>

Descripción: Encontramos nombre de algún software de dinámica de sistemas, siendo página de la universidad Politécnica de Cataluña, encontramos un contenido visto en un curso.

Sistemas Dinámicos

URL:<http://fractales.org/?q=node/82&PHPSESSID=f9179ea5e194983eb85f80952be876b5>

Descripción: Explicación de sistemas dinámicos y ejemplos de ciencias donde son útiles los sistemas dinámicos.

URL:<http://www.monografias.com/trabajos10/suaq/suaq.shtml>

Descripción: Trata de los sistemas dinámicos con probabilística, vulnerabilidad y concepto.

URL: <http://www.albany.edu/cpr/sds/>

Descripción: En esta página Web encontramos el significado de dinámica de sistemas, cuál es su uso actualmente, y define varias palabras relacionadas a este concepto para que se le haga más fácil al usuario comprender la información ofrecida. Proporciona la metodología que se utiliza en la DS, explica de dónde surgió éste campo de estudio y establece las distintas aplicaciones que han surgido desde que se dio a conocer este concepto. Establece la relación entre pensamiento sistémico y dinámica de sistemas, y explica al usuario que es la sociedad de sistemas dinámicos.

URL: http://dinamicasistemas.utalca.cl/din_sis/dinamica_sistemas.htm

Descripción: Como primer punto, éste sitio cuenta los orígenes de la Dinámica de Sistemas con la aparición del libro "Dinámica Industrial" de Jay Forrester, su fundamentación en los sistemas sociales, las limitaciones humanas y cómo es que vivimos en sistemas sin darnos cuenta de ello. Además, habla sobre la simulación, herramienta de DS, el lenguaje de diagramas que creó Forrester para el modelado y los ámbitos de las aplicaciones de los sistemas dinámicos.

Grupo de Dinámica de Sistemas

URL: <http://www.catunesco.upc.es/ads/ads.htm>

Descripción: La Universidad Politécnica de Cataluña ofrece cursos en Internet para estudiantes que tienen problemas para asistir a clases presenciales o para aquellos que estén interesados en asistir a la Universidad Politécnica de California.

Dentro del sitio existe información del "Curso de especialización de dinámica de sistemas", como la definición de "Dinámica de sistemas", aplicación de modelos de simulación, bibliotecas de modelo algunos trabajos entre otros

Capítulo Latinoamericano “Sociedad de Dinámica de Sistemas”

URL: <http://dynamica-sistemas.mty.itesm.mx/>

Descripción: Este sitio contiene información de congresos, publicaciones, cursos que se efectúan en distintos puntos en América Latina. Además de recursos relacionados con la “Dinámica de Sistemas”.

System Dynamics

URL: <http://www.sys-dyn.net/index.php>

Descripción: Ofrece servicios como automatización de procesos, red y servidores, base de datos y conversión de datos entre otros.

System Dynamics Simulation

URL: <http://www.raczynski.com/pn/simball.htm>

Descripción: El sitio contiene información de una herramienta para simulación de sistemas dinámicos llamada “SimBall”

Thinking Systemically

URL: <http://www.dpsnet.com/system/example.htm>

Descripción: Contiene un ejemplo donde aplican los 6 pasos al pensamiento sistemático, según Michael Goodman y Richard Karash

Dinámica de Sistemas

URL: <http://www.ieg.csic.es/dinamica/WEB%20Csic/dinamic/indicew.htm>

Descripción: Muestra un diagrama en el cual te explican acerca de la interrelación de los sistemas dinámicos, ejemplo de cómo construir un modelo de sistemas.

Cátedra de Dinámica de Sistemas

URL: <http://www.geocities.com/martin3162/musica.html>

Descripción: Explica el Desarrollo Sostenible de Empresas Innovadoras utilizando modelos de Dinámica de Sistemas. Hace hincapié en la problemática de las empresas y las características de la metodología de modelado; presenta gráficos sobre cada tema desarrollado, desde las diferentes etapas en la estructuración de los modelos hasta su comportamiento, etc.

Proceso de construcción de un modelo en dinámica de sistemas

URL: <http://tgs7233.galeon.com/dinamica.htm>

Descripción: Esta página desarrolla y explica de una manera bastante sencilla el proceso de construcción de un modelo en dinámica de sistemas, es decir, su metodología, conceptualización, representación o formulación, análisis y evaluación.

1.13 Resumen

Este capítulo ha presentado las definiciones, los elementos, las clasificaciones de los sistemas. Se dice que los sistemas son combinaciones por parte reunidas para obtener un resultado o formar conjunto organizados de cosas, se relaciona un todo unitario y complejo para alcanzar varios objetivos. Estos sistemas tienen como características la objetividad y la totalidad, metas o fines en los cuales se quiere llegar y los sistemas globales que tiene naturaleza orgánica.

Adicionalmente, ha introducido la necesidad de contar con una ingeniería que aplique estos conceptos en el desarrollo de mejores sistemas. Lo más importante del capítulo es que presenta el concepto de sistemas de realimentación, también conocido como sistema de ciclo cerrado o sistema dinámico complejo.

En el siguiente capítulo se presenta todo lo concerniente al control que deben tener los sistemas para poder exhibir propiedades dinámicas complejas.

1.14 Ejercicios

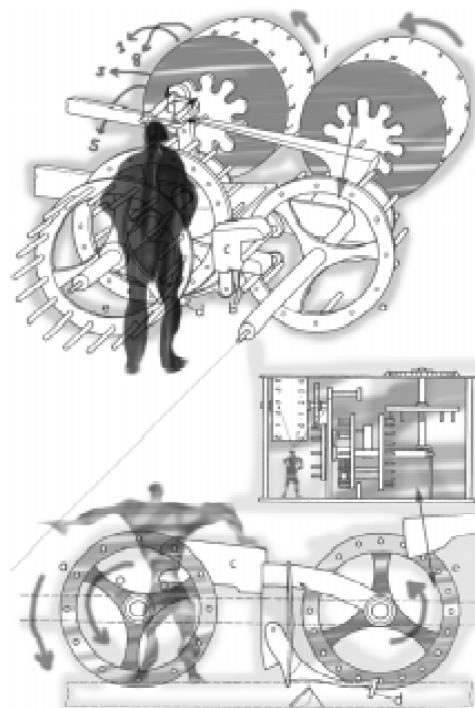
1. Mencione 5 actividades endógenas y 5 actividades exógenas de un Sistema de Transporte Urbano.
2. Defina los términos Realimentación, Sistema, Ingeniería de Sistemas.
3. Señale cuales son las etapas para diseñar un sistema dinámico.
4. Clasifique los siguientes sistemas de acuerdo a su complejidad:
 - Un carro.
 - La sonda Mars Pathfinder.
 - El árbol Panamá.
 - Los hoyos negros en el universo.
5. Clasifique los sistemas de acuerdo así son: naturales, hechos por el hombre, estático, dinámicos, adaptables, no adaptables, manuales, semiautomáticos, automáticos, abiertos, cerrados, físicos, conceptuales, de ciclo abierto, de ciclo cerrado, de realimentación positiva y negativa:

- a. Computadora.
- b. Sistema de leyes de un país.
- c. Acondicionador de Aire.
- d. Ser humano.
- e. Un reloj.
- f. El Sol.

INTRODUCCIÓN A LA TEORÍA GENERAL DE SISTEMAS

por

Oscar Johansen Bertoglio



Oscar Johansen Bertoglio

Facultad de Ciencias Económicas y Administrativas
Departamento de Administración

Título
Ingeniero Comercial, Universidad de Chile.

Magister
MBA, University of Columbia, U.S.A.

Jerarquía Académica
Profesor Titular

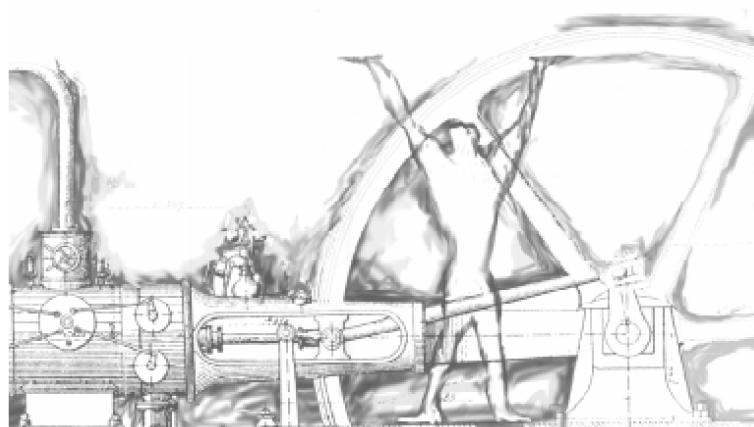
Disciplina
Administración

Especificación del Área
Administración y Sistemas.



1

Conceptos Básicos de la Teoría General de Sistemas



SISTEMA: El concepto de sistema en general está sustentado sobre el hecho de que ningún sistema puede existir aislado completamente y siempre tendrá factores externos que lo rodean y pueden afectarlo. Puleo define sistema como " un conjunto de entidades caracterizadas por ciertos atributos, que tienen relaciones entre sí y están localizadas en un cierto ambiente, de acuerdo con un cierto objetivo". También se define como un conjunto organizado de cosas o partes interactuantes e interdependientes, que se relacionan formando un todo unitario y complejo. Cabe aclarar que las cosas o partes que componen al sistema, no se refieren al campo físico (objetos), sino más bien al funcional. De este modo las cosas o partes pasan a ser funciones básicas realizadas por el sistema. Podemos enumerarlas en: entradas, procesos y salidas.

ENTIDAD: Es lo que constituye la esencia de algo y por lo tanto es un concepto básico. Las entidades pueden tener una existencia concreta, si sus atributos pueden percibirse por los sentidos y por lo tanto son medibles y una existencia abstracta si sus atributos están relacionados con cualidades inherentes o propiedades de un concepto.

ATRIBUTO: Se entiende por atributo las características y propiedades estructurales o funcionales que caracterizan las partes o componentes de un sistema

RELACION: Las relaciones internas y externas de los sistemas han tomado diversas denominaciones. Entre otras: efectos recíprocos, interrelaciones, organización, comunicaciones, flujos, prestaciones, asociaciones, intercambios, interdependencias, coherencias, etcétera. Las relaciones entre los elementos de un sistema y su ambiente son de vital importancia para la comprensión del comportamiento de sistemas vivos. Las relaciones pueden ser recíprocas (circularidad) o unidireccionales. Presentadas en un momento del sistema, las relaciones pueden ser observadas como una red estructurada bajo el esquema input/output.

SUBSISTEMA: Se entiende por subsistemas a conjuntos de elementos y relaciones que responden a estructuras y funciones especializadas dentro de un sistema mayor. En términos generales, los subsistemas tienen las mismas propiedades que los sistemas (sinergia) y su delimitación es relativa a la posición del observador de sistemas y al modelo que tenga de éstos. Desde este ángulo se puede hablar de subsistemas, sistemas o supersistemas, en tanto éstos posean las características sistémicas (sinergia).

SINERGIA: Todo sistema es sinérgico en tanto el examen de sus partes en forma aislada no puede explicar o predecir su comportamiento. La sinergia es, en consecuencia, un fenómeno que surge de las interacciones entre las partes o componentes de un sistema (conglomerado). Este concepto responde al postulado aristotélico que dice que "el todo no es igual a la suma de sus partes". La totalidad es la conservación del todo en la acción recíproca de las partes componentes (teleología). En términos menos esencialistas, podría señalarse que la sinergia es la propiedad común a todas aquellas cosas que observamos como sistemas.

FRONTERA: Los sistemas consisten en totalidades y, por lo tanto, son indivisibles como sistemas (sinergia). Poseen partes y componentes (subsistema), pero estos son otras totalidades (emergencia). En algunos sistemas sus fronteras o límites coinciden con discontinuidades estructurales entre estos y sus ambientes, pero corrientemente la demarcación de los límites sistémicos queda en manos de un observador (modelo). En términos operacionales puede decirse que la frontera del sistema es aquella línea que separa al sistema de su entorno y que define lo que le pertenece y lo que queda fuera de él (Johannsen. 1975:66).

AMBIENTE: Se refiere al área de sucesos y condiciones que influyen sobre el comportamiento de un sistema. En lo que a complejidad se refiere, nunca un sistema puede igualarse con el ambiente y seguir conservando su identidad como sistema. La

única posibilidad de relación entre un sistema y su ambiente implica que el primero debe absorber selectivamente aspectos de éste. Sin embargo, esta estrategia tiene la desventaja de especializar la selectividad del sistema respecto a su ambiente, lo que disminuye su capacidad de reacción frente a los cambios externos. Esto último incide directamente en la aparición o desaparición de sistemas abiertos.

MODELO: Los modelos son constructos diseñados por un observador que persigue identificar y mensurar relaciones sistémicas complejas. Todo sistema real tiene la posibilidad de ser representado en más de un modelo. La decisión, en este punto, depende tanto de los objetivos del modelador como de su capacidad para distinguir las relaciones relevantes con relación a tales objetivos. La esencia de la modelística sistémica es la simplificación. El metamodelo sistémico más conocido es el esquema input-output.

ELEMENTO: Se entiende por elemento de un sistema las partes o componentes que lo constituyen. Estas pueden referirse a objetos o procesos. Una vez identificados los elementos pueden ser organizados en un modelo.

ORGANIZACIÓN: N. Wiener planteó que la organización debía concebirse como "una interdependencia de las distintas partes organizadas, pero una interdependencia que tiene grados. Ciertas interdependencias internas deben ser más importantes que otras, lo cual equivale a decir que la interdependencia interna no es completa" (Buckley. 1970:127). Por lo cual la organización sistémica se refiere al patrón de relaciones que definen los estados posibles (variabilidad) para un sistema determinado.

ESTRUCTURA: Las interrelaciones más o menos estables entre las partes o componentes de un sistema, que pueden ser verificadas (identificadas) en un momento dado, constituyen la estructura del sistema. Según Buckley (1970) las clases particulares de interrelaciones más o menos estables de los componentes que se verifican en un momento dado constituyen

la estructura particular del sistema en ese momento, alcanzando de tal modo una suerte de "totalidad" dotada de cierto grado de continuidad y de limitación. En algunos casos es preferible distinguir entre una estructura primaria (referida a las relaciones internas) y una hiperestructura (referida a las relaciones externas).

INFORMACION: La información tiene un comportamiento distinto al de la energía, pues su comunicación no elimina la información del emisor o fuente. En términos formales "la cantidad de información que permanece en el sistema (...) es igual a la información que existe más la que entra, es decir, hay una agregación neta en la entrada y la salida no elimina la información del sistema" (Johannsen. 1975:78). La información es la más importante corriente negentrópica de que disponen los sistemas complejos.

CIBERNETICA: Se trata de un campo interdisciplinario que intenta abarcar el ámbito de los procesos de control y de comunicación (retroalimentación) tanto en máquinas como en seres vivos. El concepto es tomado del griego *kibernetes* que nos refiere a la acción de timonear una goleta (N.Wiener.1979).

CIRCULARIDAD: Concepto cibernético que nos refiere a los procesos de autocausación. Cuando A causa B y B causa C, pero C causa A, luego A en lo esencial es autocausado (retroalimentación, morfostasis, morfogénesis).

COMPLEJIDAD: Por un lado, indica la cantidad de elementos de un sistema (complejidad cuantitativa) y, por el otro, sus potenciales interacciones (conectividad) y el número de estados posibles que se producen a través de éstos (variedad, variabilidad). La complejidad sistémica está en directa proporción con su variedad y variabilidad, por lo tanto, es siempre una medida comparativa. Una versión más sofisticada de la TGS se funda en las nociones de diferencia de complejidad y variedad. Estos fenómenos han sido trabajados por la cibernética y están asociados a los postulados de R.Ashby (1984), en donde se

sugiere que el número de estados posibles que puede alcanzar el ambiente es prácticamente infinito. Según esto, no habría sistema capaz de igualar tal variedad, puesto que si así fuera la identidad de ese sistema se diluiría en el ambiente.

CONGLOMERADO: Cuando la suma de las partes, componentes y atributos en un conjunto es igual al todo, estamos en presencia de una totalidad desprovista de sinergia, es decir, de un conglomerado (Johannsen. 1975:31-33).

ENERGIA: La energía que se incorpora a los sistemas se comporta según la ley de la conservación de la energía, lo que quiere decir que la cantidad de energía que permanece en un sistema es igual a la suma de la energía importada menos la suma de la energía exportada (entropía, negentropía).

ENTROPIA: El segundo principio de la termodinámica establece el crecimiento de la entropía, es decir, la máxima probabilidad de los sistemas es su progresiva desorganización y, finalmente, su homogeneización con el ambiente. Los sistemas cerrados están irremediablemente condenados a la desorganización. No obstante hay sistemas que, al menos temporalmente, revierten esta tendencia al aumentar sus estados de organización (negentropía, información).

EQUIFINALIDAD: Se refiere al hecho que un sistema vivo a partir de distintas condiciones iniciales y por distintos caminos llega a un mismo estado final. El fin se refiere a la mantención de un estado de equilibrio fluyente. "Puede alcanzarse el mismo estado final, la misma meta, partiendo de diferentes condiciones iniciales y siguiendo distintos itinerarios en los procesos organísmicos" (von Bertalanffy. 1976:137). El proceso inverso se denomina **multifinalidad**, es decir, "condiciones iniciales similares pueden llevar a estados finales diferentes" (Buckley. 1970:98).

EQUILIBRIO: Los estados de equilibrios sistémicos pueden ser alcanzados en los sistemas abiertos por diversos caminos, esto

se denomina equifinalidad y multifinalidad. La mantención del equilibrio en sistemas abiertos implica necesariamente la importación de recursos provenientes del ambiente. Estos recursos pueden consistir en flujos energéticos, materiales o informativos.

EMERGENCIA: Este concepto se refiere a que la descomposición de sistemas en unidades menores avanza hasta el límite en el que surge un nuevo nivel de emergencia correspondiente a otro sistema cualitativamente diferente. E. Morin (Arnold. 1989) señaló que la emergencia de un sistema indica la posesión de cualidades y atributos que no se sustentan en las partes aisladas y que, por otro lado, los elementos o partes de un sistema actualizan propiedades y cualidades que sólo son posibles en el contexto de un sistema dado. Esto significa que las propiedades inmanentes de los componentes sistémicos no pueden aclarar su emergencia.

FUNCION: Se denomina función al output de un sistema que está dirigido a la mantención del sistema mayor en el que se encuentra inscrito.

HOMEOSTASIS: Este concepto está especialmente referido a los organismos vivos en tanto sistemas adaptables. Los procesos homeostáticos operan ante variaciones de las condiciones del ambiente, corresponden a las compensaciones internas al sistema que sustituyen, bloquean o complementan estos cambios con el objeto de mantener invariante la estructura sistémica, es decir, hacia la conservación de su forma. La mantención de formas dinámicas o trayectorias se denomina **homeorrosis** (sistemas cibernéticos).

INPUT / OUTPUT: Los conceptos de input y output nos aproximan instrumentalmente al problema de las fronteras y límites en sistemas abiertos. Se dice que los sistemas que operan bajo esta modalidad son procesadores de entradas y elaboradores de salidas.

Input.- Todo sistema abierto requiere de recursos de su ambiente. Se denomina input a la importación de los recursos (energía, materia, información) que se requieren para dar inicio al ciclo de actividades del sistema.

Output.- Se denomina así a las corrientes de salidas de un sistema. Los outputs pueden diferenciarse según su destino en servicios, funciones y retroinputs.

PROCESO: El proceso es lo que transforma una entrada en salida, como tal puede ser una máquina, un individuo, una computadora, un producto químico, una tarea realizada por un miembro de la organización, etc. En la transformación de entradas en salidas debemos saber siempre como se efectúa esa transformación. Con frecuencia el procesador puede ser diseñado por el administrador. En tal caso, este proceso se denomina "Caja blanca". No obstante, en la mayor parte de las situaciones no se conoce en sus detalles el proceso mediante el cual las entradas se transforman en salidas, porque esta transformación es demasiado compleja. Diferentes combinaciones de entradas o su combinación en diferentes órdenes de secuencia pueden originar diferentes situaciones de salida. En tal caso la función de proceso se denomina una "caja negra".

CAJA NEGRA: La caja negra se utiliza para representar a los sistemas cuando no sabemos que elementos o cosas componen al sistema o proceso, pero sabemos que a determinadas corresponden determinadas salidas y con ello poder inducir, presumiendo que a determinados estímulos, las variables funcionaran en cierto sentido.

MORFOGENESIS: Los sistemas complejos (humanos, sociales y culturales) se caracterizan por sus capacidades para elaborar o modificar sus formas con el objeto de conservarse viables (retroalimentación positiva). Se trata de procesos que apuntan al desarrollo, crecimiento o cambio en la forma, estructura y estado del sistema. Ejemplo de ello son los procesos de diferenciación,

la especialización, el aprendizaje y otros. En términos cibernéticos, los procesos causales mutuos (circularidad) que aumentan la desviación son denominados morfogenéticos. Estos procesos activan y potencian la posibilidad de adaptación de los sistemas a ambientes en cambio.

MORFOSTASIS: Son los procesos de intercambio con el ambiente que tienden a preservar o mantener una forma, una organización o un estado dado de un sistema (equilibrio, homeostasis, retroalimentación negativa). Procesos de este tipo son característicos de los sistemas vivos. En una perspectiva cibernética, la morfostasis nos remite a los procesos causales mutuos que reducen o controlan las desviaciones.

NEGENTROPIA: Los sistemas vivos son capaces de conservar estados de organización improbables (entropía). Este fenómeno aparentemente contradictorio se explica porque los sistemas abiertos pueden importar energía extra para mantener sus estados estables de organización e incluso desarrollar niveles más altos de improbabilidad. La negentropía, entonces, se refiere a la energía que el sistema importa del ambiente para mantener su organización y sobrevivir (Johannsen. 1975).

RECURSIVIDAD: Proceso que hace referencia a la introducción de los resultados de las operaciones de un sistema en él mismo (retroalimentación).

RETROALIMENTACION: Son los procesos mediante los cuales un sistema abierto recoge información sobre los efectos de sus decisiones internas en el medio, información que actúa sobre las decisiones (acciones) sucesivas. La retroalimentación puede ser negativa (cuando prima el control) o positiva (cuando prima la amplificación de las desviaciones). Mediante los mecanismos de retroalimentación, los sistemas regulan sus comportamientos de acuerdo a sus efectos reales y no a programas de outputs fijos. En los sistemas complejos están combinados ambos tipos de corrientes (circularidad, homeostasis).

Retroalimentación negativa.- Este concepto está asociado a los procesos de autorregulación u homeostáticos. Los sistemas con retroalimentación negativa se caracterizan por la mantención de determinados objetivos. En los sistemas mecánicos los objetivos quedan instalados por un sistema externo (el hombre u otra máquina).

Retroalimentación positiva.- Indica una cadena cerrada de relaciones causales en donde la variación de uno de sus componentes se propaga en otros componentes del sistema, reforzando la variación inicial y propiciando un comportamiento sistémico caracterizado por un autorreforzamiento de las variaciones (circularidad, morfogénesis). La retroalimentación positiva está asociada a los fenómenos de crecimiento y diferenciación. Cuando se mantiene un sistema y se modifican sus metas/fines nos encontramos ante un caso de retroalimentación positiva. En estos casos se aplica la relación desviación-amplificación (Mayurama. 1963).

RETROINPUT: Se refiere a las salidas del sistema que van dirigidas al mismo sistema (retroalimentación). En los sistemas humanos y sociales éstos corresponden a los procesos de autorreflexión.

SERVICIO: Son los outputs de un sistema que van a servir de inputs a otros sistemas o subsistemas equivalentes.

SISTEMAS ABIERTOS: Se trata de sistemas que importan y procesan elementos (energía, materia, información) de sus ambientes y esta es una característica propia de todos los sistemas vivos. Que un sistema sea abierto significa que establece intercambios permanentes con su ambiente, intercambios que determinan su equilibrio, capacidad reproductiva o continuidad, es decir, su viabilidad (entropía negativa, teleología, morfogénesis, equifinalidad).

SISTEMAS CERRADOS: Un sistema es cerrado cuando ningún elemento de afuera entra y ninguno sale fuera del sistema. Estos

alcanzan su estado máximo de equilibrio al igualarse con el medio (entropía, equilibrio). En ocasiones el término sistema cerrado es también aplicado a sistemas que se comportan de una manera fija, rítmica o sin variaciones, como sería el caso de los circuitos cerrados.

SISTEMAS CIBERNETICOS: Son aquellos que disponen de dispositivos internos de autocomando (autorregulación) que reaccionan ante informaciones de cambios en el ambiente, elaborando respuestas variables que contribuyen al cumplimiento de los fines instalados en el sistema (retroalimentación, homeorrosis).

SISTEMAS TRIVIALES: Son sistemas con comportamientos altamente predecibles. Responden con un mismo output cuando reciben el input correspondiente, es decir, no modifican su comportamiento con la experiencia.

TELEOLOGIA: Este concepto expresa un modo de explicación basado en causas finales. Aristóteles y los Escolásticos son considerados como teleológicos en oposición a las causalistas o mecanicistas.

VARIABILIDAD: Indica el máximo de relaciones (hipotéticamente) posibles ($n!$).

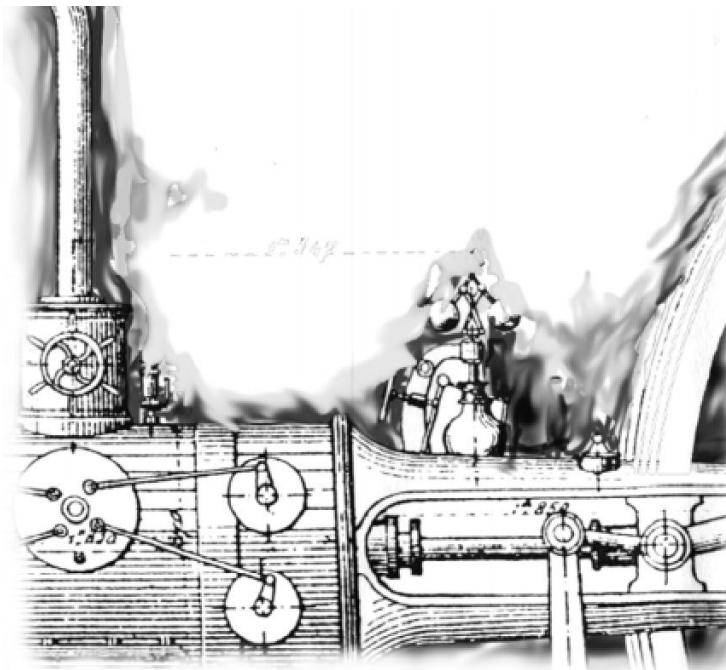
VARIEDAD: Comprende el número de elementos discretos en un sistema (v = cantidad de elementos).

VIABILIDAD: Indica una medida de la capacidad de sobrevivencia y adaptación (morfostásis, morfogénesis) de un sistema a un medio en cambio.

INTRODUCCIÓN

2

Introducción a la Teoría General de Sistemas



Al enfoque de sistemas puede llamársele correctamente teoría general de sistemas aplicada. Es importante proporcionar una comprensión básica del surgimiento de la ciencia de los sistemas generales.

Delinearemos las principales propiedades de los sistemas y de los dominios de sistemas. Además, se hace una comparación entre los supuestos subyacentes a los enfoques analítico-mecánicos. Esta comparación demuestra la incapacidad de los enfoques para tratar el dominio de los campos biológico, conductual, social y similares.

TEORIA GENERAL DE SISTEMAS

La teoría de la organización y la práctica administrativa han experimentado cambios sustanciales en años recientes. La información proporcionada por las ciencias de la administración y la conducta ha enriquecido a la teoría tradicional. Estos esfuerzos de investigación y de conceptualización a veces han llevado a descubrimientos divergentes. Sin embargo, surgió un enfoque que puede servir como base para lograr la convergencia, el enfoque de sistemas, que facilita la unificación de muchos campos del conocimiento. Dicho enfoque ha sido usado por las ciencias físicas, biológicas y sociales, como marco de referencia para la integración de la teoría organizacional moderna. El primer expositor de la Teoría General de los Sistemas fue Ludwig von Bertalanffy, en el intento de lograr una metodología integradora para el tratamiento de problemas científicos. La meta de la Teoría General de los Sistemas no es buscar analogías entre las ciencias, sino tratar de evitar la superficialidad científica que ha estancado a las ciencias. Para ello emplea como instrumento, modelos utilizables y transferibles entre varios continentes científicos, toda vez que dicha extrapolación sea posible e integrable a las respectivas disciplinas. La Teoría General de los Sistemas se basa en dos pilares básicos: aportes semánticos y aportes metodológicos:

APORTES SISTEMÁTICOS: Las sucesivas especializaciones de las ciencias obligan a la creación de nuevas palabras, estas se acumulan durante sucesivas especializaciones, llegando a formar casi un verdadero lenguaje que sólo es manejado por los especialistas.

De esta forma surgen problemas al tratarse de proyectos interdisciplinarios, ya que los participantes del proyecto son especialistas de diferentes ramas de la ciencia y cada uno de ellos maneja una semántica diferente a los demás.

APORTES METODOLOGICOS:

Jerarquía de los Sistemas:

Al considerar los distintos tipos de sistemas del universo Kennet Boulding proporciona una clasificación útil de los sistemas donde establece los siguientes niveles jerárquicos:

- 1. Primer nivel, Estructura Estática:** Se le puede llamar nivel de los marcos de referencia.
- 2. Segundo nivel, Sistema Dinámico Simple:** Considera movimientos necesarios y predeterminados. Se puede denominar reloj de trabajo.
- 3. Tercer nivel, Sistema Cibernético:** El sistema se autorregula para mantener su equilibrio.
- 4. Cuarto nivel, Sistema Abierto:** En este nivel se comienza a diferenciar la vida. Puede de considerarse nivel de célula.
- 5. Quinto nivel, Genético-Social:** Está caracterizado por las plantas.
- 6. Sexto nivel, Sistema Animal:** Se caracteriza por su creciente movilidad, comportamiento teleológico y su autoconciencia.

- 7. Séptimo nivel, Sistema Humano:** Es el nivel del ser individual, considerado como un sistema con conciencia y habilidad para utilizar el lenguaje y símbolos.
- 8. Octavo nivel, Sistema Social o Sistema de Organizaciones Humanas:** Considera el contenido y significado de mensajes, la naturaleza y dimensiones del sistema de valores, la transcripción de imágenes en registros históricos, sútiles simbolizaciones artísticas, música, poesía y la compleja gama de emociones humanas.
- 9. Noveno nivel, Sistemas Trascendentales:** Completan los niveles de clasificación: estos son los últimos y absolutos, los ineludibles y desconocidos, los cuales también presentan estructuras sistemáticas e interrelaciones.

DESARROLLO HISTÓRICO DE LA TEORÍA DE SISTEMAS

La primera formulación en tal sentido es atribuible al biólogo Ludwig von Bertalanffy (1901-1972), quien acuñó la denominación "Teoría General de Sistemas". Para él, la TGS debería constituirse en un mecanismo de integración entre las ciencias naturales y sociales y ser al mismo tiempo un instrumento básico para la formación y preparación de científicos. Sobre estas bases se constituyó en 1954 la Society for General Systems Research, cuyos objetivos fueron los siguientes:

- a. Investigar el isomorfismo de conceptos, leyes y modelos en varios campos y facilitar las transferencias entre aquellos.
- b. Promoción y desarrollo de modelos teóricos en campos que carecen de ellos.
- c. Reducir la duplicación de los esfuerzos teóricos
- d. Promover la unidad de la ciencia a través de principios conceptuales y metodológicos unificadores.

Como ha sido señalado en otros trabajos, la perspectiva de la TGS surge en respuesta al agotamiento e inaplicabilidad de los enfoques analítico-reduccionistas y sus principios mecánico-causales (Arnold & Rodríguez, 1990b). Se desprende que el principio clave en que se basa la TGS es la noción de totalidad orgánica, mientras que el paradigma anterior estaba fundado en una imagen inorgánica del mundo. A poco andar, la TGS concitó un gran interés y pronto se desarrollaron bajo su alero diversas tendencias, entre las que destacan la cibernetica (N. Wiener), la teoría de la información (C. Shannon y W. Weaver) y la dinámica de sistemas (J. Forrester). Si bien el campo de aplicaciones de la TGS no reconoce limitaciones, al usarla en fenómenos humanos, sociales y culturales se advierte que sus raíces están en el área de los sistemas naturales (organismos) y en el de los sistemas artificiales (máquinas). Mientras más equivalencias reconozcamos entre organismos, máquinas, hombres y formas de organización social, mayores serán las posibilidades para aplicar correctamente el enfoque de la TGS, pero mientras más experimentemos los atributos que caracterizan lo humano, lo social y lo cultural y sus correspondientes sistemas, quedarán en evidencia sus inadecuaciones y deficiencias.

No obstante sus limitaciones, y si bien reconocemos que la TGS aporta en la actualidad sólo aspectos parciales para una moderna Teoría General de Sistemas Sociales (TGSS), resulta interesante examinarla con detalle. Entendemos que es en ella donde se fijan las distinciones conceptuales fundantes que han facilitado el camino para la introducción de su perspectiva, especialmente en los estudios ecológico culturales (e.g. M. Sahlins, R. Rappaport), politológicos (e.g. K. Deutsch, D. Easton), organizaciones y empresas (e.g. D. Katz y R. Kahn) y otras especialidades antropológicas y sociológicas. Finalmente, el autor quiere agradecer a Juan Enrique Opazo, Andrea García, Alejandra Sánchez, Carolina Oliva y Francisco Osorio, quienes dieron origen a este documento en una versión de 1991, bajo el proyecto de investigación SPITZE.

PAPEL DE LA TEORÍA GENERAL DE SISTEMAS

Esta teoría se ha desarrollado con la finalidad de ofrecer una alternativa a los esquemas conceptuales conocidos con el nombre de enfoque analítico y mecánico con la aplicación del método científico. Se les llama mecánico porque estos fueron instrumentos en el desarrollo de las leyes de Newton, y analítico estos proceden por medio del análisis, se caracterizan porque pueden ir de lo más complejo a lo más simple.

Los enfoques analíticos y mecánicos sufrieron las siguientes omisiones:

1. Estos no podían explicar por completo, los fenómenos como organización, mantenimiento, regulación y otros procesos biológicos.
2. El método analítico no fue adecuado para el estudio de los sistemas que tuvieron que ser tratados holísticamente, las propiedades del sistema de esta clase no podían inferirse de las propiedades de las partes, un supuesto importante del enfoque analítico y mecánico.
3. Las teorías mecánicas no fueron diseñadas para tratar con sistemas de complejidad organizada, ya que estas mostraban estructuras más complejas acopladas a fuertes interacciones.

La teoría general de sistema ha evolucionado para ofrecer un marco de trabajo conceptual y dialéctico en el cual pueden desarrollarse los métodos científicos adecuados a otros sistemas y no propiamente a los del mundo físico, y pueden lograr:

1. Adoptan un enfoque holístico hacia los sistemas.
2. Provocan la generalidad de leyes particulares, mediante el hallazgo de similitudes de estructura (isomorfismo) a través de los sistemas.
3. Anima el uso de modelos matemáticos, cambian el énfasis de una consideración de contenido a una

- estructura, la cual ayuda en la solución de muchas controversias de utilidad cuestionable.
4. Promueve la unida de la ciencia, al proporcionar un marco de referencia coherente para la organización del conocimiento.

TEORÍA GENERAL DE SISTEMAS Y LA UNIDAD DE LA CIENCIA

A la par de las matemáticas y la filosofía con la cual se pregunta por la unidad de la ciencia, el hombre ha desarrollado modelos para estudiar y comprender las relaciones de las estructuras y los fenómenos del mundo real, los cuales pueden tomar distintas formas, pero ellos están hechos para lograr una mejor comprensión de la complejidad del mundo real. Estos complejos surgen en dos niveles diferentes: el micronivel, que se interesa por las relaciones básicas de causa y efecto, estas regulan el desempeño de los componentes elementales; y el macronivel, es en donde se estudian las interrelaciones entre los subsistemas elementales.

LA TEORÍA GENERAL DE SISTEMAS Y LA REALIDAD

La Teoría General de Sistemas describe un nivel de construcción teórico de modelos que se sitúa entre las construcciones altamente generalizadas de las matemáticas puras y las teorías específicas de las disciplinas especializadas y que en estos últimos altos ha hecho sentir, cada vez más fuerte, la necesidad de un cuerpo sistemático de construcciones teóricas que pueda discutir, analizar y explicar las relaciones generales del mundo empírico. Según Boulding ese es el destino de la Teoría General de Sistemas. Por supuesto que no se busca establecer una teoría general de prácticamente cualquier cosa, única y total, que reemplace todas las teorías especiales de cada disciplina en particular.

Tal teoría, en la práctica, no tendría contenido, porque en la medida que aumentamos la generalidad tenemos que hacerlo a costa del contenido. Por ejemplo, se puede pensar en una persona en particular. Sin embargo, podemos generalizarla diciendo que es un ciudadano de una ciudad determinada. Hemos ganado en generalización, pero hemos perdido en cuanto al contenido particular de la persona. Pero podemos llegar fácilmente a un segundo grado de generalización diciendo que es un hombre de una determinada nacionalidad. Luego podemos generalizarlo más aún, pensando en su sentido genérico: es un sistema vivo, y aún más, en otro grado de generalización es un sistema natural, por fin podemos decir que es un sistema abierto y, más aún, un sistema y finalmente un objeto.

Sin embargo en alguna parte, entre lo específico que no tiene significado y lo general que no tiene contenido, debe existir para cada propósito y para cada nivel de abstracción, un grado óptimo de generalidad. Los teóricos de sistemas afirman que este óptimo grado de generalidad en teoría no siempre es alcanzado por las ciencias en particular

Este punto de vista se ve cada vez más demostrado o adquiere mayor fuerza, cuando uno contempla las nuevas disciplinas que se crean y que representan, fundamentalmente, la “tierra de nadie” que separa a las disciplinas concretas. Así hablamos de físico-química (que no es ni física pura ni química pura), de psicología social (que no es ni psicología pura ni sociología pura) y, más reciente aún, de bioquímica, biofísicoquímica (y no sería extraño que ya se pensara en términos de psicobiofísicoquímica o sociopsicobiofísicoquímica). En este sentido, la teoría de sistemas (o el enfoque de sistemas) toma una posición contraria (como metodología) al enfoque reduccionista que discutimos anteriormente. Mientras ese último tiende a la subdivisión cada vez mayor del todo, y al estudio particular de esas subdivisiones, el enfoque de sistemas pretende integrar las partes hasta alcanzar una totalidad lógica o de una independencia o autonomía relativa con respecto a la totalidad mayor de la cual también forma parte.

EL ENFOQUE DE LOS SISTEMAS

ENFOQUE REDUCCIONISTA:

Ejemplo 1: Hace un tiempo atrás, mientras me preparaba a efectuar un viaje fuera del país, tuve que ir al consultorio del medico, a quien visito periódicamente por una enfermedad crónica, con el fin de que me diera un certificado, explicando mi enfermedad, y que podría servirme como un antecedente en previsión de alguna afección que pudiera sufrir mientras estuviera en el extranjero.

Mientras esperaba al médico gastroenterólogo observó en una de las paredes de su clínica un gran cuadro que representaba las diferentes partes del organismo, cada una dentro de los contornos de la figura humana. Así, la primera figura representaba el esqueleto; la segunda, el aparato circulatorio; la tercera, el sistema digestivo, la cuarta el sistema muscular y la quinta, el sistema nervioso.

Cada una de ellas mostraba una parte de la anatomía humana, separada de tal modo que facilitara su estudio y la comprensión de las funciones de cada sistema en particular. Sin embargo, superponiéndolas de cierta manera se llegaba a ser humano como tal.

Es evidente que es a través de esas divisiones como la biología ha logrado estudiar e investigar la anatomía humana. Es decir, el progreso alcanzado por estas ciencias se debe, en gran parte, a lo que, generalmente, se denomina el enfoque reduccionista, en el cual se estudia un fenómeno complejo a través del análisis de sus elementos o partes componentes.

Ejemplo 2: Observemos un problema trivial. A pesar de que muchos partidos de fútbol importantes son televisados, normalmente podemos observar el estadio lleno y la reventa de entradas, es decir, una fuerte presión para ver el juego desde allí. ¿Es que esa gente no dispone de un receptor de TV o le es

imposible "visitar" a algún familiar o amigo que lo tenga? Creemos que no. Ver un partido de fútbol en TV, dice el aficionado, "no es lo mismo que verlo en la cancha". Aparte del ingrediente marginal (pero importante) del ruido, del contacto entre los espectadores, en fin, del estado emocional que provoca una contienda de equipos importantes, es difícil seguir el juego desde la pantalla del televisor. Uno observa al arquero efectuar un rechazo, ¿hacia dónde? No lo sabemos, hasta que la pantalla, siguiendo la trayectoria del balón, nos indica hacia qué jugador o posición éste iba dirigido. Lo mismo ocurre en casi todo el partido, excepto en los pases cortos. ¿Qué sucede? Simplemente, que la actual tecnología no nos permite "observar" toda la cancha desde la pantalla de TV. Solo nos muestra el lugar donde se desarrolla la acción central (donde está el balón en juego), pero no nos permite observar el todo, el cuadro general, los movimientos de los jugadores sin el balón, los desplazamientos y las demarcaciones. En una palabra, observamos una parte del conjunto que no nos permite "gozar" del espectáculo completo.

Ejemplo 3: Para dejar más clara la idea, y utilizando la imaginación del lector, supongamos que pudiéramos disponer de un aparato tal que nos permitiera observar solamente la conducta de un determinado jugador de uno de los dos equipos que se enfrentan en el estadio. Aparece sólo el individuo en acción. Evidentemente que, al cabo de unos minutos, nos parecería que este hombre se conduce de una manera bastante extraña que nosotros no comprendemos: le vemos correr, detenerse, saltar, caer al suelo, levantar las manos, moverse con un comportamiento errático. Sin embargo, si en un momento dado apretamos un botón de nuestro televisor y lo integramos al comportamiento del resto de los jugadores, árbitros y público, entonces comprenderemos y nos explicaremos cabalmente una conducta hasta entonces extraña y absurda.

TEORÍA DE LOS CAMPOS: Ya en los años 30, Kurt Lewin, el famoso psicólogo fundador de la escuela basada en la “**teoría de los campos**” (*Field Theory*) para el estudio del comportamiento humano y de grupos señalaba que “lo que resulta importante en la teoría del campo es la forma en que procede el análisis. En vez de escoger uno u otro elemento aislado dentro de una situación, la importancia del cual no puede ser juzgada sin tomar en cuenta la situación como un todo, **la teoría del campo encuentra ventajoso, como regla, comenzar por la caracterización de las situación como un todo**”. Despues de la primera aproximación, los diversos aspectos y partes de la situación son sometidos a un análisis cada vez mas específico y detallado. Es obvio que este método es la mejor manera para no errar el camino, engañados por uno u otro elemento de la situación”. Sin duda que Lewin pensaba ya en la idea integracionista, porque se enfrentaba a un objeto de estudio: el hombre y/o los grupos, que son sistemas bastante más complejos que un pedazo de mineral o una célula.

LA TEORÍA GENERAL DE SISTEMAS Y SU APLICACIÓN

CIBERNÉTICA: Esta nueva ciencia, desarrollada por Norbert Wiener en su clásico libro “Cibernetica”,¹⁰ se basa en el principio de la retroalimentación (o causalidad circular) y de homeostasis; explica los mecanismos de comunicación y control en las maquinas y los seres vivos que ayudan a comprender los comportamiento generados por estos sistemas que se caracterizan por sus propósitos, motivados por la búsqueda de algún objetivo, con capacidades de auto - organización y de auto - control.

Según S. Beer, Wiener, al definir la cibernetica como la “ciencia de la comunicación y el control en el animal y en la maquina”, apuntaba a las leyes de los sistemas complejos que permanecen invariables cuando se transforma su materia. Considerándola en su sentido más amplio, Beer la define como “la ciencia de la organización efectiva”. Allí señala que las leyes de los sistemas complejos son invariables, no frente a las transformaciones de su

materia, sino también de su *contenido*. Nada importa, dice Beer, que el contenido del sistema sea neurofisiológico, automotor, social o económico.

TEORÍA DE LA INFORMACIÓN: Esta introduce el concepto de información como una cantidad mensurable, mediante una expresión isomórfica con la entropía negativa en física. En efecto, los matemáticos que han desarrollado esta teoría han llegado a la sorprendente conclusión de que la fórmula de la información es exactamente igual a la fórmula de la entropía, sólo con el signo cambiado, de donde se deduce que:

$$\begin{aligned} \text{Información} &= - \text{entropía} \\ \text{Información} &= \text{neguentropía} \end{aligned}$$

Ahora bien la entropía (positiva en física es una medida de desorden. Luego la información (o entropía negativa) o neguentropía es una medida de organización. En este sentido, es interesante observar una conclusión a que ha llegado J.J. Miller que señala que, mientras más complejos son los sistemas (entendiéndose por complejidad el número posible de estados que puede presentar cada parte y el número de las posibles relaciones entre esas partes) mayor es la energía que dichos sistemas destinan tanto a la obtención de la información como a su procesamiento, decisión, almacenaje y/o comunicación.

TEORÍA DE LOS JUEGOS: Desarrollada por Morgenstern y, principalmente, por von Neuman, trata de analizar, mediante un novedoso marco de referencia matemática, la competencia que se produce entre dos o más sistemas racionales (o por parte de un sistema) antagonista, los que buscan maximizar sus ganancias y minimizar sus pérdidas (es decir, buscan alcanzar o “jugar” la estrategia óptima).

A través de esta técnica se puede estudiar el comportamiento de partes en conflicto, sean ellas individuos, oligopolios o naciones. Evidentemente, aun los supuestos sobre los cuales descansa esta teoría son bastante restrictivos (suponen conducta racional

entre los competidores), sin embargo, su avance, es decir, la eliminación, c, al menos, la extensión o mayor flexibilidad de los supuestos dependerá del avance realizado no sólo en este campo, sino en campos afines, como son la conducta o dinámica de grupos y, en general, la o las teorías que tratan de explicar y resolver (o predecir) los conflictos.

TEORÍA DE LA DECISIONES: En general, en este campo se han seguido dos líneas diferentes de análisis. Una es la Teoría de la Decisión misma que busca analizar, en una forma parecida a la Teoría de los juegos, la selección racional de alternativas dentro de las organizaciones o sistemas sociales. Se basa en el examen de un gran número de situaciones y sus posibles consecuencias, determinando así (por procedimientos estadísticos, fundamentalmente basados en la toma de las probabilidades), una decisión que optimice el resultado

La otra línea de análisis, encabezada básicamente por H.A. Simón, es el estudio de la "conducta" que sigue el sistema social, en su totalidad y en cada una de sus partes, al afrontar el proceso de decisiones. Esto ha conducido a una teoría "conductista" de la empresa a diferencia de la teoría económica, muy en boga entre los economistas que han desarrollado la teoría de la competencia perfecta y/o imperfecta (Boulding Chamberling, y otros). En ella se estudia el comportamiento de estos sistemas sociales que se caracterizan por perseguir ciertos objetivos.

Esta aproximación ha modificado sustancialmente la teoría administrativa al describir el comportamiento de los centros de decisiones, enfatizando el problema de las comunicaciones y sus riesgos, etc.

TOPOLOGÍA O MATEMÁTICA RELACIONAL: La Topología ha sido reconocida como un área particular de las matemáticas en los últimos 50 años, y su principal crecimiento se ha originado dentro de los últimos 30 años. Es una de las nuevas ramas de las matemáticas que ha demostrado, más poder y ha producido fuertes repercusiones en la mayoría de las antiguas ramas de

esta ciencia y ha tenido también efecto importante en las otras ciencias, incluso en las ciencias sociales. Partió como una respuesta a la necesidad del análisis clásico del cálculo y de las ecuaciones diferenciales. Sin embargo, la topología no es una rama del análisis, sino una especie de geometría, una geometría más bien de pensamiento geométrico basado en la prueba de la existencia de un cierto teorema, en campos tales como las redes, los gráficos, los conjuntos.

Su aplicación al estudio de las interacciones entre las partes de los sistemas (sociales o de otro tipo) se hace evidente. Por ejemplo, L. Spier expresa la teoría de los gráficos como un método, para comprender la conducta administrativa. Señala que es una gran ayuda para ilustrar las propiedades estructurales de un problema administrativo, o de una estructura organizacional y las propiedades de las conexiones entre sus partes.

EL ANÁLISIS FACTORIAL: Es decir él aislamiento, por medio del análisis matemático, de los factores en aquellos problemas caracterizados por ser multivariados.

Su aplicación se ha concentrado en diferentes áreas; dentro de las ciencias sociales especialmente en psicología.

En esta ciencia, este planteamiento trata de determinar las principales dimensiones de los grupos (por ejemplo, en el estudio de la dinámica de grupos), mediante la identificación de sus elementos claves. Esto significa que se puede medir en un gran grupo una cantidad de atributos y determinar un número bastante más limitado de dimensiones independientes, por medio de las cuales pueda ser más económico y funcionalmente definido medir cualquier grupo particular de una población grupal mayor. En la dinámica de grupos se define como "sintalidad" 10 que el término de *personalidad* define en el individuo. Los factores principales encontrados por los psicólogos sociales que apoyan este enfoque son los de energía, habilidad y dirección.

INGENIERÍA DE SISTEMAS: Se refiere a la planeación, diseño, evaluación y construcción científica de sistemas hombre - maquina. El interés teórico de este campo se encuentra en el hecho de que aquellas entidades cuyos componentes son heterogéneos (hombres, maquinas, edificios, dinero y otros objetos, flujos de materias primas, flujos de producción, etc.) pueden ser analizados como sistemas o se les puede aplicar el análisis de sistemas.

La Ingeniería de sistemas de acuerdo con Hall es una parte de la técnica creativa organizada que se ha desarrollado como una forma de estudiar los sistemas complejos (especialmente industriales). El aumento de la complejidad se pone de manifiesto con el creciente número de interacciones entre los miembros de una población en crecimiento, la acelerada división del trabajo y la especialización de las funciones, el empleo creciente de las maquinas que reemplazan a la mano de obra, con el consiguiente aumento de la productividad y la creciente velocidad y volumen en las comunicaciones y transporte.

INVESTIGACIÓN DE OPERACIONES: Es el control científico de los sistemas existentes de hombres, maquinas, materiales, dinero, etc. Quizás la definición más moderna avanzada en este campo sea la de Stafford Beer, uno de los primeros participantes en el Operational Research, que se creó en Inglaterra durante la Segunda Guerra Mundial, y que, formado por sabios y técnicos de las diferentes ramas del saber, se enfrentó y resolvió problemas particulares presentados por las fuerzas armadas.

Beer define a la investigación de operaciones como: "El ataque de la ciencia moderna a los complejos problemas que surgen de la dirección y la administración de los grandes sistemas compuestos por hombres, maquinas, materiales y dinero en la industria, el comercio, el gobierno y la defensa. Su enfoque distintivo es el desarrollo de un modelo científico del sistema incorporando factores tales como el azar y el riesgo, con los cuales predecir y comparar los resultados de las diferentes decisiones, estrategias o controles alternativos. El propósito es ayudar a la administración

a determinar su política y sus acciones de una manera científica". Esta definición después de muchas consultas con los principales expertos británicos en este campo fue adoptada por la "Operational Research Society of Great Britain".

DEFINICIONES DE SISTEMAS

3

Los Sistemas del Mundo Real



La definición del sistema (o el establecimiento de sus fronteras) puede no ser un problema simple de resolver. Es posible hacer varios intentos de definición hasta que por fin encontramos una que encierre nuestra unidad de análisis y sus principales interrelaciones con el medio (incluyendo aquellas fuerzas de su medio que pueden modificar, y de hecho modifican la conducta de esa unidad de análisis).

La dificultad de fijar las fronteras de los sistemas se debe a las siguientes características de éstos:

1. Es bastante difícil (si no imposible) aislar los aspectos estrictamente mecánicos de un sistema. Por ejemplo, al escribir estas líneas, puedo pensar y estoy viendo como mi mano y sus dedos aprisionan el lápiz y con ciertos movimientos determinados se deslizan sobre el papel. Sin embargo, mal podría explicar este lápiz fenómeno si me dedico a observar sólo el sistema mano - lápiz - papel. Indudablemente debe agregar el sistema molecular y las actividades neurales y/o los procesos interpretativos del cerebro.
2. El intercambio o la relación entre sistemas no se limita exclusivamente a una familia de sistemas. Existe un contacto permanente con el mundo exterior. Para escribir estas líneas, mi sistema no sólo está formado por brazo, cerebro, lápiz y papel, sino además por un conjunto de libros y apuntes desparramados sobre el escritorio y que sirven de apoyo a mi trabajo. Existe aquí un continuo cambio de energía y de información entre mi sistema y el mundo exterior.
3. Finalmente existe un continuo intercambio de interrelaciones tiempo - secuencia, pensamos que cada efecto tiene su causa, de modo que las presiones del medio sobre el sistema modifican su conducta y, a la vez, este cambio de conducta modifica al medio y su comportamiento. Las opiniones de cierto autor modifican mis ideas sobre algún aspecto de la

materia que estoy escribiendo, pero podría ser que lo que finalmente escriba modifique las ideas de ese autor.

En todo caso, para la definición de un sistema siempre contaremos con dos conceptos que pueden ser de gran ayuda: la idea de un supersistema y la idea de los subsistemas. De este modo, podemos definir a nuestro sistema en relación con su medio inmediato, por una parte, y en relación con sus principales componentes, por otra.

Así, si mi interés es estudiar una playa de arena, bien puedo limitar mi vista a esa playa, y su frontera estará dada por sus límites geográficos. Pero a su vez, podría definir el supersistema como los objetos que se encuentran inmediatamente al otro lado de las fronteras del sistema (parte del mar y el continente) y que, a mi juicio, inciden fundamentalmente en la conducta del sistema. Por otro lado, puedo definir los subsistemas, que podrían ser en este caso el grano de arena, las rocas, etc. y su constitución o características. Sin duda que, al tomar estos tres niveles de organización para estudiar el nivel del medio, **estaremos asegurándonos una mejor comprensión del comportamiento del nivel intermedio de organización que es, precisamente, el que deseamos estudiar.**

COMPONENTES: Los componentes de un sistema son simplemente las partes identificables del mismo.

ATRIBUTOS DE LOS COMPONENTES: Los componentes, por ser objetos o personas, poseen propiedades o características. Estas influyen en la operación del sistema, en su velocidad, precisión, confiabilidad, capacidad y muchos otros aspectos. Un ejemplo de la elección entre máquinas pudiera ser la selección de un dispositivo de salida entre varias opciones: despliegue visual, un sistema de audio, una impresora mecánica o un graficador. En el sistema de información, las características del componente de salida no son una alta velocidad sino la claridad, la economía y la permanencia relativa. Por tanto, se escoge una impresora como auxiliar de la computadora, y el formato de la

salida es un informe "impreso" que se envía a la gerencia una vez al mes.

ESTRUCTURA: La estructura de un sistema es el conjunto de relaciones entre los objetos y atributos de los objetos de un sistema. El grado en que los elementos funcionan juntos para alcanzar los objetivos totales sirve para definir la estructura. Los elementos de un sistema trabajan juntos en grados variables.

- 1 Relaciones Disfuncionales.**- Estas pueden presentarse a raíz de fenómenos naturales, atributos antagónicos o conflictos organizacionales. Por ejemplo, los departamentos de producción y mercadotecnia tal vez no coincidan en los programas de producción, en los inventarlos o en el servicio a clientes.
- 2 Relaciones Parasitarias.**- Son aquellas en que uno aprovecha los recursos del otro sin dar nada a cambio. Por ejemplo, en Estados Unidos hay algunas comisiones gubernamentales cuyos miembros reciben un sueldo y lo único que hacen es reunirse una vez por año.
- 3 Relaciones Simbióticas.**- Se presentan entre organismos u organizaciones diferentes que satisfacen mutuamente sus necesidades. Por ejemplo, el garrapatero que vive de las vacas y que devora garrapatas, con lo cual satisface la necesidad de eliminarlas del cuerpo de ese animal. He aquí otro ejemplo: una compañía y sus vendedores se necesitan mutuamente y colaboran en una relación simbiótica.
- 4 Relaciones Sinergéticas.**- Estas relaciones existen a veces entre los subsistemas y los elementos; se refuerzan entre sí en la obtención de objetivos comunes.
- 5 Relaciones Optimizadas.**- En ellas los intercambios de recursos y los intercambios de objetivos de los subsistemas mantienen un equilibrio dinámico para optimizar la salida total del sistema en forma continua. Se trata con toda seguridad de

un sistema ideal, mientras que el sistema sinergético que lucha por la optimización representa a los sistemas reales.

INTERFASES: La interfaz es una conexión entre dos sistemas, la región de contacto. La interfaz entre el hombre y la computadora es la salida que corresponde a la entrada de la computadora. Eso pudiera ser la preparación de tarjetas perforadas o la terminal. La computadora

ENTROPÍA: La entropía es el movimiento de un sistema hacia un desgaste, desorden o discrepancia totales. Un sistema cerrado alcanza su entropía máxima cuando se descompone. En los sistemas biológicos o sociales, la entropía puede ser invertida por las entradas de información y energía, llegando con ello a mayores estados de orden y organización. A esto se le llama incremento de la entropía negativa.

CLASIFICACION DE LOS SISTEMAS

SISTEMAS NATURALES Y ARTIFICIALES: Los sistemas naturales abundan en la naturaleza. La ecología de la vida es un sistema natural, y cada organismo es un sistema natural especial. El sistema del agua del mundo, por lo menos antes que el hombre lo modificara, era un sistema, como también es el sistema solar. Sus objetivos varían enormemente. Un sistema se centrara en la defensa nacional; otro Será un sistema de transporte. La organización de una campaña es un sistema con muchos sistemas más pequeños incorporados a él (producción, contabilidad, etc.) y otros, como los sistemas de comunicaciones y los de distribución de oficinas, sobrepuerto a la principal organización económica de la gente.

SISTEMAS SOCIALES. HOMBRE-MAQUINA Y MECÁNICOS: Distinguiéndose de otros sistemas, objetivos y procesos. Las empresas, las dependencias gubernamentales, los partidos políticos, los clubes sociales y las sociedades técnicas son

ejemplos de sistemas que pueden estudiarse desde esta perspectiva.

La mayor parte de los sistemas empíricos caen dentro de la categoría de hombre – maquina. En la actualidad, casi todos los hombres emplean equipos de una u otra clase en sus trabajos organizados

Los sistemas puramente mecánicos deben obtener sus propias entradas y mantenerlas.

Los sistemas mecánicos totalmente autosuficientes y autorreparables todavía pertenecen a la ciencia ficción, pese a que algunos sistemas eléctricos generadores de energía eléctrica se acercan cada vez más a la autosuficiencia

SISTEMAS ABIERTOS Y CERRADOS: El sistema abierto es aquel que interactúa con su ambiente. Todos los sistemas que contienen organismos vivos son abiertos, porque en ellos influye lo que es percibido por los organismos. En un sentido más importante, as organizaciones suelen ser sistemas que operan dentro de otros más extensos y, por lo mismo son abiertos. Por ejemplo, la estructura de mercadotecnia de una empresa es un sistema que forma parte de otro más grande: la compañía entera. Y ésta a su vez es un sistema en el interior del sistema industrial global.

El hecho de que una compañía interactúe con su ambiente (un sistema más amplio) hace de ella un sistema abierto

El sistema ambiental con el que mas se relaciona es el sistema de la industria de la cual forman parte.

Prosiguiendo en este orden de ideas, señalamos que la industria es parte del sistema económico del país, el cual a su vez es un sistema dentro de la sociedad, la sociedad es un sistema en el interior del sistema mundial y este forma parte forma parte del sistema solar; y así sucesivamente hasta llegar a lo desconocido

Es más difícil entender de qué cosa constituye un sistema cerrado. El ambiente que rodee a un sistema cerrado no cambia y, si lo hace, se levantara una barrera entre el ambiente y él para impedir cualquier influencia. Aunque es poco probable que

existen realmente los sistemas cerrados, este concepto tiene importantes implicaciones

Cuando emprendemos experimentos en el laboratorio para estudiar el comportamiento humano, estamos intentando establecer temporalmente un sistema cerrado.

En el mundo de los negocios los problemas son resueltos algunas veces como si hubiera un sistema cerrado; ello se hace con el propósito de simplificar la situación lo suficiente para que se obtenga, por lo menos, una primera aproximación.

SISTEMAS PERMANENTES Y TEMPORALES: Relativamente pocos sistemas artificiales son permanentes. Sin embargo, en la práctica se dice que son "permanentes" aquellos que duran mucho más que las operaciones que en ellos realiza el ser humano. Nuestro sistema económico, que está cambiando gradualmente, es esencialmente permanente respecto a nuestros planes para el futuro.

Los sistemas verdaderamente temporales están destinados a durar cierto periodo y luego desaparecen.

Un proyecto pequeño de una investigación en grupo realizada en el-laboratorio es un sistema temporal.

SISTEMAS ESTABLES Y NO ESTABLES: Un sistema estable es aquel cuyas propiedades y operaciones no varían de manera importante o lo hacen sólo en ciclos repetitivos. La fábrica automatizada, la dependencia gubernamental que procesa los pagos al seguro social, los planteles de enseñanza media y el sistema de transbordador son ejemplos de este tipo de sistema. Una empresa publicitaria, un sistema de defensa continental, un laboratorio de investigación y desarrollo, un ser humano son ejemplos de sistemas no estables

SUBSISTEMAS Y SUPRASISTEMAS: En los apartados anteriores se ha advertido que cada sistema está incluido en un sistema mas grande. El sistema en la jerarquía que más nos interesa estudiar o controlar suele llamarse "el sistema". La empresa mercantil se considera como "el sistema" o "el sistema total", cuando el interés se centra en la producción, en la distribución de los bienes y en las fuentes de utilidades e

ingresos. Como dice Stanford L. Optner: "El sistema total consta de todos los objetos, atributos y relaciones necesarias para alcanzar los objetivos dadas varias restricciones". La palabra sistema se utiliza con mucha frecuencia en el sentido de sistema total. El objetivo de este último define la finalidad para la cual todos los objetos, atributos y relaciones del sistema han sido organizados.

Los sistemas más pequeños incorporados al sistema reciben el nombre de subsistemas. Esta distinción tiene importantes implicaciones en la práctica respecto a la optimización y al enfoque de sistemas", según veremos más adelante.

El suprasistema denota sistemas extremadamente grandes y complejos. El suprasistema puede referirse a cualquier sistema que incluya al que está estudiándose. La economía puede considerarse un suprasistema en relación con la empresa mercantil.

SISTEMAS ADAPTATIVOS Y NO ADAPTATIVOS: Otra gama o espectro de posibilidades del sistema abarca los Adaptativos y lo no Adaptativos.

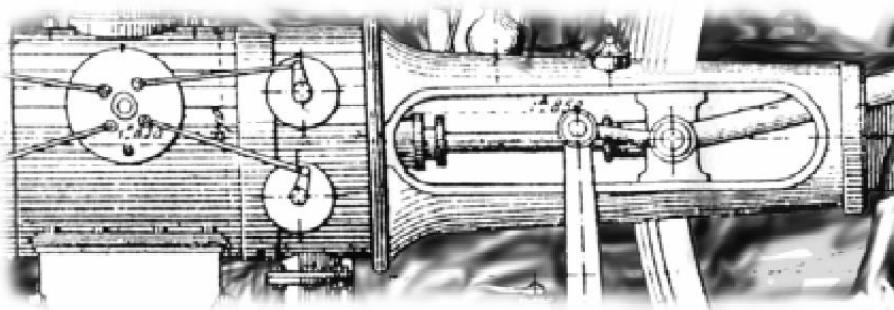
Un sistema que reacciona con su ambiente en tal forma que mejora su funcionamiento, logro o probabilidad de supervivencia se llama sistema adaptativo. Los organismos vivos de alto nivel, entre ellos los animales y el hombre, se sirven de la adaptación para afrontar las amenazas de los cambios en el ambiente físico o los que se producen en su sociedad. La teoría evolucionista se basa sobre todo en el concepto de un sistema adaptativo.

Por ultimo asociamos a la adaptación la fuente de energía, el aprendizaje y la modificación de si mismo. Por ejemplo, si las computadoras pudieran conectarse a una fuente de energía duradera, "aprender" a modificarse y repararse sin intervención del hombre, se convertirían en sistemas adaptivos

EL MUNDO EN EQUILIBRIO

4

El Principio de la Organicidad



El mundo puede ser representado como un sistema o como una colección de sistemas o sub sistemas que actúan y se interrelacionan unos con otros dentro de una realidad dinámica.

A pesar de toda esta enorme dinámica de fuerzas, de acción y reacción entre los diferentes sistemas, no existe un caos, si no un cierto orden y equilibrio que dan mas una impresión de acciones y relaciones suaves mas que pronunciadas.

Este fenómeno de acción equilibrada puede ser explicado a partir de dos concepciones diferentes.

Una de ellas es el aparente equilibrio del sistema según la mecánica newtoniana y la otra es la teoría general de sistemas.

EL EQUILIBRIO Y LAS LEYES DE NEWTON

Isaac Newton (1642 - 1727) definió varias leyes sobre el movimiento o mecánica.

La primera de ellas señala que cada objeto o cuerpo persiste en un estado de descanso o inmóvil, o con un movimiento uniforme en línea recta, hasta que sea forzado a cambiar de este estado por fuerzas ejercidas contra él.

En el caso de los sistemas, sabemos que estos están compuestos de millones de sub sistemas que no parecen estar inmóviles, ¿entonces, como podríamos explicar esta contradicción aparente con la ley de newton?; pues para esto newton presento su tercera ley que dice: a cada acción sigue una reacción igual, vale decir que la acción de mutua de dos cuerpos, del uno sobre el otro es siempre igual y en dirección opuesta.

Esta tercera ley de newton a dado origen al principio de acción - reacción que señala que cada acción se encuentra acompañada de una reacción, también se conoce este principio como el de causa - efecto.

Entonces podemos decir que el mundo permanece en equilibrio gracias a que los cambios que se producen entre los subsistemas se cancelan unos con otros, permaneciendo así el sistema total (la tierra) en equilibrio.

EL EQUILIBRIO Y LA TEORIA GENERAL DE SISTEMAS

Desde el punto de vista de la variabilidad del sistema total en relación a la variabilidad de sus partes existe una definición que expresado en palabras significa que **un objeto es un sistema cuando la variabilidad que experimenta la totalidad es menor que la suma de las variabilidades de cada una de sus partes o componentes.**

Este fenómeno lo podemos observar, en las actividades de la totalidad de las abejas en un panal.

Mirando en forma global, nos parece estar observando una enorme variabilidad en la conducta particular de cada abeja o grupo de abejas. Sin embargo, si consideramos al panal como una totalidad, podremos observar que su conducta es bastante equilibrada y ello nos permite predecir el comportamiento del panal como una empresa productora de miel de abejas.

Si unimos los dos conceptos, la variabilidad de los subsistemas y la variabilidad del medio, podemos comprender el equilibrio que puede mostrar un sistema.

Frente a los cambios externos que se producen en el medio, el sistema provisto de los homeostatos (derivado de homeostasis) necesarios, aminorá esos impactos, desarrollando programas pre establecidos que tienden a hacer posible una serie de reacciones internas del sistema que lo defienden de las variaciones del medio.

Por otra parte, la sinergia tiende a nivelar los cambios internos que sufren los subsistemas.

Todo esto hace que el sistema tenga la propiedad de autocontrol y de autorregulación que lo lleva hacia un equilibrio homeostático o hacia un “estado permanente” que se caracteriza por la mantención de una relación determinada y estable entre la energía que entra al sistema (corriente de entrada) y la energía que sale del sistema (corriente de salida).

LA EVOLUCION EN EQUILIBRIO

En el corto plazo, las acciones y reacciones de un sistema no aparecen reflejadas en el carácter general del sistema, pero son fuerzas latentes que tratan de llevar al sistema a un cambio muchas veces impredecible.

Existen entonces dos fuerzas o dos aspectos fundamentales en el comportamiento de los sistemas.

Uno de ellos son las fuerzas que resisten los cambios bruscos y severos, y el otro aspecto es que los ciclos son rara vez o nunca similares.

En otras palabras, existen en la naturaleza fuerzas que buscan mantener un tipo particular de equilibrio al resistir los cambios rápidos, y fuerzas que demandan cambios, pero producidos por procesos lentos y evolutivos.

EL PRINCIPIO DE LA ORGANICIDAD

Se denomina principio de organicidad al proceso de evolución que tiende a aumentar el grado de organización que poseen los sistemas (sistemas abiertos y en especial los sistema vivos).

Parece ser que existe una tendencia natural, inherente a los sistemas vivos hacia la organización.

Y esta tendencia en muchos casos es independiente de los centros “ejecutivos” o directrices de esos sistemas.

La teoría de la evolución habla de una organización cada vez más compleja y la segunda ley, se refiere a la entropía creciente.

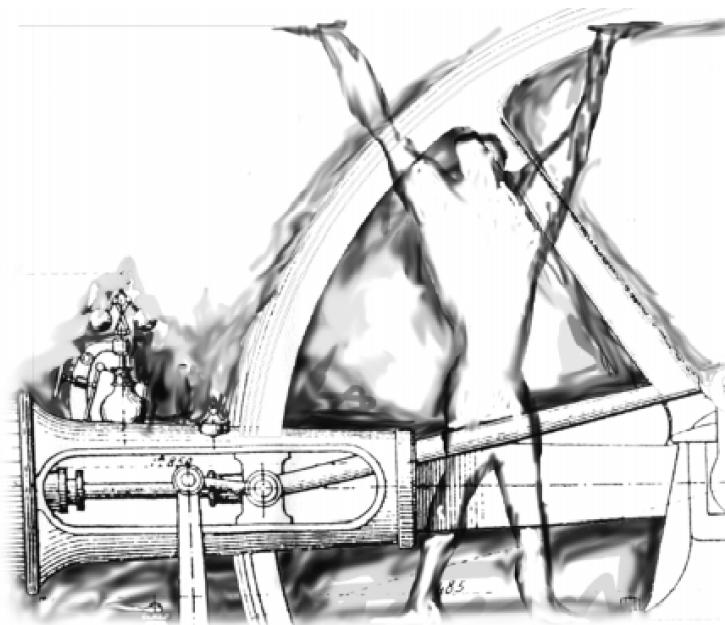
Estos dos principios esbozan una contradicción más aparente que real, es decir que se puede salir del dilema planteado.

Ya al discutir la entropía, Wolfgang Wiese planteaba la opinión de que, junto a las conocidas leyes de la energía, debería existir una tercera ley, la ley de la organización.

Señala que la organización de un sistema es un principio que no se puede referir a fuerza o materia, “pero que, por si, es una magnitud independiente, ni energía ni sustancia, si no algo tercero expresado por la medida y el modo de orden”.

5

Definición de un Sistema Total



En el desarrollo de la teoría y el enfoque de sistemas, nos hemos referido a los sistemas en general (sistema social, mecánico, humano). Ahora, tomando en cuenta específicamente a los sistemas sociales, y más concretamente, a aquéllos cuyo objetivo es proporcionar bienes y/o servicios a la comunidad, es decir, en la empresa. Nuestro objetivo será discutir la definición o identificación de un sistema desde el punto de vista de un investigador que desea analizarlo.

Supongamos que el investigador se encuentra en condiciones de observar el todo o **sistema total** y tratemos de determinar los pasos que dará para alcanzar el punto desde el cual podrá observar ese todo. Su método, como lo plantea Churchman, "consiste en definir cuidadosamente de qué está hablando".

El término "sistema"; aunque puede ser definido de muchas, se le define como: "El conjunto de partes coordinadas para alcanzar ciertos objetivos".

Para esto, debemos explicar qué es lo que entendemos por "partes coordinadas".

Específicamente, el objetivo del investigador de sistemas es definir cuidadosamente y en detalle cuál es el sistema total, el medio en que se encuentra, cuáles son sus objetivos y sus partes y cómo estas partes apoyan el logro de esos objetivos.

El investigador de sistemas deberá seguir los siguientes pasos si quiere su objetivo: Describir y Definir un **sistema total**.

Los pasos son los siguientes:

- 1. Los objetivos del sistema total;
- 2. El medio en que vive el sistema;
- 3. Los recursos del sistema;
- 4. Los componentes del sistema;
- 5. La dirección del sistema;

Sin embargo, en ningún caso, estos pasos deben, forzosamente, tomarse de acuerdo con la secuencia aquí presentada. Más bien,

a medida que uno avanza en el análisis y descripción del sistema, es probable que uno deba reexaminar el trabajo realizado en los pasos previos.

Este es un proceso lógico y la lógica es esencialmente un proceso de controlar y recontrolar nuestros razonamientos.

Con estas advertencias en mente, discutiremos estas cinco consideraciones básicas que el investigador debe tener presentes cuando se enfrenta con la tarea de definir y describir un sistema.

LOS OBJETIVOS DEL SISTEMA TOTAL

La definición de los objetivos de un sistema total no siempre es tarea fácil. Puede existir confusión en su determinación. Generalmente los participantes del sistema no se preocupan, aun cuando sus objetivos y definiciones puedan tener una serie de propósitos independientes de la actuación del sistema.

Churchman presenta el siguiente ejemplo:

El director de un departamento universitario desea obtener el mayor presupuesto posible para la operación de su unidad. Como consecuencia, debe defender su posición en diferentes comités, comisiones o consejos que forman parte de la estructura de la universidad.

En estas ocasiones debe plantear los objetivos del departamento en la forma más atractiva posible, ya que debe buscar el máximo prestigio y el máximo poder "político" con el fin de obtener ese mayor presupuesto.

Señalará la importancia de la docencia en ese campo, el valor que tienen para la comunidad los proyectos de investigación que en él se desarrollan, la alta calificación académica que poseen sus integrantes, etc.

Estos son algunos de los argumentos que plantea y que menciona como los objetivos del departamento en cuestión.

En forma similar, el gerente de una empresa (a través de su departamento de relaciones públicas) o el administrador de una

agencia estatal, deben presentar una imagen social a través de los objetivos de esos sistemas. No lo hacen sólo para atraer clientela, sino también para atraer capitales y apoyo del medio. Recuérdese que la aceptación o "legalización" del sistema en el medio es una condición necesaria para su supervivencia.

En general, estas declaraciones de principios son denominadas **objetivos** por los miembros del sistema, pero para nuestros propósitos de investigación y análisis ellas son demasiado vagas y, en muchos casos, equivocadas.

Este tipo de objetivos ha sido denominado por algunos estudiosos como "no operacionales" y equivalen, como señalábamos hace poco, más bien a declaraciones de principios o propósitos que a objetivos concretos sobre los cuales dirigir la conducta del sistema.

Si tomamos estos objetivos no operacionales en forma demasiado seria, podemos llegar a errores y equivocaciones en la identificación de los objetivos reales del sistema comparado con aquellos definidos por ese sistema.

El gerente de una empresa industrial puede hacernos pensar que el objetivo de su empresa es prestar servicio a la comunidad, sobre la base de alguna utilidad, etcétera.

Estos objetivos no operacionales tampoco pueden dejarse totalmente de lado. En efecto, ellos cumplen una función bastante específica en la empresa o en el sistema social de que se trate:

Son objetivos generalmente no conflictivos y, por lo tanto, pasan a ser un mecanismo de cohesión del grupo humano que forma el sistema.

¿Qué investigador de un departamento universitario no va a estar de acuerdo en que él está allí para aumentar el conocimiento en ese campo determinado del saber?

¿Qué ejecutivo medio o alto de una empresa va a estar contra la idea de los "objetivos sociales" de la empresa?

Una forma, sugerida por Churchman, en que el investigador puede determinar los objetivos del sistema, es observando si el sistema, a sabiendas, sacrificará otros objetivos con tal de alcanzar aquellos definidos por autoridades del sistema.

Por supuesto que, así como es difícil definir los verdaderos objetivos de una persona, así también puede serlo determinar los verdaderos objetivos de un sistema.

Como personas, generalmente ocultamos nuestros verdaderos objetivos, pues nos parece que pueden no ser satisfactorios desde el punto de vista de otras personas o de la comunidad en que vivimos. Si ellos fueran publicitados, es probable que tuviéramos problemas en obtener los diferentes tipos de apoyos que necesitamos en la vida (empleo, aceptación social, etc.).

El propósito, entonces, del investigador es determinar aquellos objetivos verdaderos y operacionales.

Operacionales en el sentido que pueden ser medidos y que a través de esta medición se pueda determinar la calidad de la actuación del sistema, o la forma como está operando éste.

Algunos estudiosos de las organizaciones industriales señalan que se debe hacer una distinción entre los **objetivos reales y los objetivos legitimados** del sistema.

Estos últimos tienen que ver con la moral o la ética de los objetivos.

Por ejemplo, un investigador, siguiendo la filosofía de la teoría de sistemas puede definir los objetivos de un sistema de carreteras en términos de la cantidad de vehículos que pueden pasar a través de segmentos específicos de una de ellas dentro de un período dado de tiempo. Sin embargo, puede que este objetivo no se encuentre "legitimado" desde el punto de vista social, no sólo por el costo de los accidentes, sino por el ruido, por ciertos problemas ecológicos y los inconvenientes que pueden surgir para los pueblos y aldeas por donde dicha carretera pasa (y para aquellos que quedan fuera de ella).

Por lo tanto, en esta tarea de describir los objetivos reales del sistema (a través de la medición de su actuación) el investigador debe ser persistente (aunque con ello eventualmente pueda entrar en conflicto con sus colegas o clientes) y estar siempre alerta para no caer en la "falacia de los objetivos evidentes".

EL MEDIO DEL SISTEMA

Una vez que el investigador ha logrado clasificar los objetivos del sistema (o la medición de su actuación) el aspecto siguiente que debe estudiar y considerar es el medio que lo rodea. Este puede ser definido como aquello que está fuera, que no pertenece al sistema, que se encuentra más allá de sus "fronteras". También puede ser ésta una tarea difícil, pues no siempre es sencillo lograr este resultado.

Si observamos un automóvil, uno puede pensar, en un primer momento, que el medio de este sistema es todo aquello que está fuera del automóvil.

Incluso podemos decir que todo lo que está más allá de la pintura exterior del vehículo conforma su medio.

¿Pero esto es correcto?

¿Es correcto afirmar que lo que queda más allá, o fuera de las paredes de una fábrica es el medio de ese sistema?

La fábrica puede tener representantes en diversos puntos del país, y aún en el extranjero, ya sea para la venta de sus productos o para la compra de equipos y materiales. Estas son, sin duda, partes del sistema total que constituye esa empresa industrial, y, sin embargo, estas partes no se encuentran dentro de sus paredes. Para complicar más este caso, es posible que el gerente general de esa empresa pertenezca a un determinado grupo de poder, a través del cual pueda ejercer ciertas presiones políticas y así obtener determinadas ventajas para esa empresa. Sus actividades políticas podrían ser consideradas como pertenecientes al sistema, aunque difícilmente podrían ocurrir dentro del espacio físico ocupado por la empresa. Esto nos puede llevar a concluir que, posiblemente, la pintura exterior del utomóvil

no sea el límite o la frontera de ese sistema, como no lo es en el caso de la fábrica.

Se ha dicho que en esta edad de la tecnología eléctrica, el teléfono ha llegado a ser prácticamente una parte del individuo humano. En muchos casos parece difícil establecer una diferencia entre el oído. No podemos eliminar el teléfono que ayuda al oído. No podemos eliminar el teléfono, así como no podemos eliminar el oído de la persona. Así considerado, el teléfono sería una parte del sistema que hemos denominado ser humano.

Por lo tanto, el investigador de sistemas debe tener un criterio sobre el medio que se encuentre más allá de la observación de sus fronteras aparentes. Un criterio para enfrentar este problema es considerar que, cuando señalamos que algo queda fuera del sistema, queremos indicar que el sistema prácticamente no tiene control sobre ello, es decir, poco o nada puede hacer para modificar sus características o su conducta.

El medio corresponde a los "datos dado" al sistema y, evidentemente, desde este punto de vista constituye sus limitaciones.

El medio de un sistema estará determinado por el problema que tiene entre manos el investigador y, evidentemente, una forma de determinarlo es fijando las fronteras reales del sistema *de acuerdo con el problema concreto*.

El medio no es sólo aquello que se encuentra fuera del control del sistema, sino que también es algo que determina, en parte, la conducta de éste.

Sin duda alguna que uno de los aspectos más importantes del medio de un sistema social es su "clientela", o la demanda. Por supuesto que el sistema puede influir en ella a través de la publicidad, de los precios y en general, de la comercialización de su producto (sea éste un bien o un servicio). Pero, en la medida que la demanda sea determinada por el individuo externo, se

encontrará fuera del sistema, en su medio y, pasará a constituir un dato o un factor limitante para ese sistema.

Un buen método para determinar si un aspecto determinado pertenece al medio o al sistema, lo proporciona C.W. Churchman. Señala que el medio no es el aire que respiramos, el grupo social al que pertenecemos o a la casa en que vivimos, no importa cuánto estos elementos parecieran estar fuera de nosotros. En cada caso uno debe hacerse dos preguntas:

- ¿Puedo hacer algo frente a ello?
- ¿Tiene importancia para mis objetivos?

Si la primera pregunta tiene una respuesta negativa y la segunda una positiva, ese aspecto constituye nuestro medio.

LOS RECURSOS DEL SISTEMA

Cuando hablamos de los recursos del sistema nos estamos refiriendo a *su interior*, es decir, a sus recursos internos. Por lo tanto no deben ser confundidos con los recursos externos; es decir, aquellas fuentes de energía o de información que llegan al sistema a través de sus corrientes de entrada. Estos se encuentran fuera del sistema, pertenecen al medio.

Los recursos del sistema son los arbitrios de que dispone para llevar a cabo el proceso de conversión y para mantener la estructura interna; en una palabra, para sobrevivir.

En realidad, existen ciertos recursos que pueden ser considerados tanto como recursos externos y como recursos internos.

Por ejemplo, la fuerza laboral de un sistema. Evidentemente que un ejecutivo, al estudiar la implementación de un proyecto o una decisión puede decir: cuento con tantos recursos humanos y, en ese sentido, estará hablando de recursos internos.

Pero cuando se enfrenta a una huelga general o de todo ese sector industrial o a un paro en la locomoción, puede decir que su corriente de entrada de energía humana se ha detenido o ha decrecido, de acuerdo con el caso. A nuestro juicio, la diferencia se encuentra en determinar si esa fuerza laboral pertenece al sistema o al medio y, de acuerdo nuevamente con el problema, aunque se trate de un mismo grupo de trabajadores, en un momento serán parte del sistema, como en el caso de la implementación del proyecto, y en otro serán parte del medio, como el caso de la huelga.

En general, los recursos del sistema, como opuestos al medio, son todo aquello que el sistema puede cambiar o utilizar para su propia ventaja. El sistema puede decidir cuáles trabajadores harán eso y cuáles aquello, o dónde se invertirá esta parte del presupuesto y en qué se gastará esa otra, o qué equipo y cuánto tiempo se utilizará en una determinada operación. En una palabra, son recursos internos del sistema aquellos sobre los cuales éste posee control.

Por eso, cuando nos referimos a los recursos, generalmente lo hacemos en términos de dinero, horas-hombre y equipos.

Así como a veces era difícil definir los objetivos reales del sistema o su medio, también puede suceder lo mismo al definir los recursos con que éste cuenta. Ya observamos alguna dificultad en el caso de la mano de obra o fuerza laboral que puede ser recurso interno o externo de acuerdo al problema que se trate.

Una forma de observar y determinar los recursos de un sistema social (especialmente de una empresa) es a través de su **balance general**. Efectivamente, en él aparecen diferentes recursos que el sistema posee; por ejemplo: sus edificios, equipos, cuentas por cobrar, saldos en bancos y caja, inventarios de materias primas, repuestos, productos terminados y otros. Es decir, todos aquellos recursos cuyo valor puede ser transformado en dinero. Sin

embargo, si nos detenemos aquí, sin duda alguna sólo habríamos contabilizado parte de los recursos de ese sistema.

En efecto, hay recursos que no pueden ser convertidos tan fácilmente en dinero y otros que por una u otra razón no aparecen en el balance general.

Tomemos, por ejemplo, el personal con que cuenta el sistema. Sin duda alguna que son recursos del sistema su nivel de educación, su experiencia, sus grados académicos y/o profesionales; su cantidad, su distribución o promedio de edad, determinan mejores o peores recursos humanos.

Otro caso se encuentra en la "imagen" que el sistema posee en el medio. Sin duda alguna que este es un recurso que la empresa puede explotar en su beneficio.

Incluso puede pensarse que son recursos propios del sistema el poder o la influencia que algunos de sus integrantes pueda tener en otros sistemas.

Por ejemplo, si por alguna relación particular resulta que uno de los ejecutivos de una empresa es muy amigo (o familiar) de un alto administrador de un banco comercial, esta relación puede ser considerada como un recurso del sistema, ya que a través de este canal, el sistema puede incluso ejercer algún control sobre esa oficina, quitándosela así al medio y sumándola al sistema. El sistema puede obtener ciertas ventajas tales como descuento de letras, préstamos a corto o a largo plazo, etcétera.

Existe, sin embargo, otra objeción aún más seria a la simple utilización del balance general. En efecto, este balance nos muestra *como fueron* utilizados los recursos. El investigador de sistemas debe aprender de la experiencia de la historia pasada. Pero el balance general típico oculta casi toda la información importante que uno desearía conocer si desea aprender de la experiencia. Lo que realmente debe servir como elemento de enseñanza son las oportunidades perdidas o mal aprovechadas, las posibilidades que nunca lograron implementarse porque los recursos estaban ocupados en otras materias y quizás empleados en forma ineficiente (es decir, observar los diferentes "costos de

sustitución"). Sin embargo, estas informaciones no se encuentran en ese documento.

Desde este punto de vista, se hace necesario el establecimiento de todo un sistema de comunicación e información para proveer a los ejecutivos con los conocimientos necesarios para sus decisiones. Y, en este sentido, un sistema de información y comunicación eficiente es un valioso recurso.

Un sistema social puede tener recursos reales y también recursos potenciales. Estos últimos es probable que no sean considerados o que nunca se hagan reales, porque normalmente se requiere de una cierta inversión para que puedan cambiar de estado. Pero el hecho es que, ya sea a través del avance tecnológico o por otra razón, es posible que los recursos actuales del sistema puedan ser altamente amplificados.

El mismo desarrollo de un sistema de información y De comunicación (posiblemente utilizando computadores u otros equipos tecnológicos sofisticados) puede aumentar la capacidad y la calidad de las decisiones de los ejecutivos incrementando sus recursos administrativos, o liberándolos para otras actividades más productivas. En ese sentido, una buena organización y una buena delegación de decisiones rutinarias por parte de los ejecutivos, pueden darle a éstos las oportunidades de atacar problemas de real importancia que de otro modo no habrían sido tocados o en cuyo análisis y solución habrían tenido que invertirse otros recursos.

Por estas razones, para la contabilización de los recursos propios del sistema, no sólo es necesario incluir los recursos reales, sino considerar también los recursos potenciales y las posibilidades que existen para transformarlos en reales, pues los mecanismos o componentes que sirven para aumentar o amplificar los recursos del sistema pueden ser los más importantes.

LOS COMPONENTES DEL SISTEMA

Los recursos propios forman la reserva general del sistema a partir de la cual se puede desarrollar su conducta para alcanzar sus objetivos reales. Las acciones específicas que se llevan a cabo las realizan sus componentes, sus partes o sus subsistemas.

¿Cuáles son los subsistemas? Una forma de determinarlos podría ser a través del organismo que muestra las diferentes unidades administrativas (siempre que se dé en ellos el principio de la recursividad) en que se ha dividido el sistema, tomando así las ventajas de la división del trabajo o diferenciación y de la especialización. Así tenemos las divisiones, los departamentos, las secciones, etc. Sin embargo, un cuidadoso examen del sistema puede indicarnos que esos pueden no ser los componentes *reales*, aunque posean títulos que en un principio nos llevarán a tomarlos como tales.

Por ejemplo, en las empresas industriales podemos encontrar fácilmente un departamento de producción, lo que nos hace pensar que es allí donde se desarrolla toda la función convertidora del sistema. Otro departamento puede denominarse comercialización, y esto también nos puede llevar a pensar que es allí donde se dirige y administra la corriente de salida. Ahora, supongamos que nos encontramos en una empresa que fabrica según pedido (es decir, que el producto se hace de acuerdo a las características solicitadas por los clientes). En este caso ¿no debe estar presente, en el momento de la venta, el equipo de producción para aceptar o no las peticiones del cliente en función de las capacidades y recursos con que cuenta el subsistema? Aquí, indudablemente la venta depende de producción; luego, es posible que ambos subsistemas sean en realidad uno solo. Por estas razones, la estructura organizacional tradicional con sus funciones de comercialización, producción, adquisiciones, financieras, relaciones industriales etc. no es una buena ayuda para el investigador de sistemas que pretende definir las partes

de ese sistema, quien debe utilizar otro criterio. Este puede ser, de acuerdo con Churchman, observar las "misiones", "los trabajos" o "las actividades" básicas, es decir, la división racional de las tareas que el sistema debe llevar a cabo. Así, por ejemplo, para la municipalidad de una ciudad, sus actividades básicas consisten en limpieza y sanidad, cuidado y reparación de calles y veredas, controles de tránsito y del comercio, salud y seguridad pública, etc. Por lo tanto, un investigador de sistemas interesado en estudiar las partes del sistema municipal, podrá enumerar estas misiones, trabajos o actividades y enseguida colocar bajo cada una de ellas las diferentes oficinas, agencias, organismos e instituciones, sin preocuparse de su nombre (salvo que todos son municipales) y definir así las partes y sus integrantes. Es posible que un organismo forme parte de más de un subsistema.

Por ejemplo, la fuerza de carabineros es parte, sin duda alguna, del subsistema encargado de la seguridad pública. Sin embargo, ese mismo organismo participa en el control del tráfico. Si un analista de sistemas desea ubicar y definir el subsistema educacional dentro del sistema de gobierno de un país, comprenderá rápidamente que esta función no sólo se lleva a cabo dentro de los límites de un ministerio de educación.

También se encuentran comprometidos en actividades educacionales otros departamentos (y ministerios) a través del entrenamiento y capacitación de su personal y del público en general, ya sea a través de folletos, cursos cortos, programas de TV, etc. Por lo tanto, la evaluación social de las tareas o actividades educacionales no puede efectuarse solamente dentro de las líneas departamentales tradicionales.

En esta tarea de identificar las partes o componentes del todo, es probable que el analista encuentre problemas serios, especialmente con aquellas personas que dirigen los departamentos o unidades administrativas del sistema. El director de una unidad de la línea tradicional, tiene claramente definidos sus sistemas particulares y lo distingue del todo; debe sostener batallas en defensa de su unidad, ya sea en términos de

presupuestos o en términos de personal; aún más, su trabajo o su participación es evaluada normalmente en términos de la conducta de su departamento. Y esto no sucede sólo con el jefe del departamento, sino también con las personas que participan dentro de él. Existe generalmente una identificación entre los *individuos* y su unidad de trabajo y aún es posible que, a través del desarrollo de las organizaciones informales, estos departamentos solidifiquen sus fronteras, aislándose del exterior, es decir, de los otros componentes del sistema

Por ejemplo, en las universidades esto sucede frecuentemente. El estudio de las matemáticas aparece hoy día, prácticamente en todos los programas de estudio de las diferentes carreras; sin embargo, es el departamento de matemáticas el que define lo que estas materias significan, la forma en que deben impartirse y los niveles que deben alcanzarse en cada caso.

En general las unidades administrativas del sistema buscan la máxima independencia por razones de orgullo, poder, status, etcétera.

¿Por qué debe ser tan persistente el investigador de sistemas en hablar de misión en vez de departamento?

Simplemente porque al hablar de misión y analizarla, él puede estimar el valor de una actividad para el sistema total, lo que no es posible lograr a través de la estimación de la realización de un departamento (o de su valor). El debe saber si la actividad de un componente del sistema es mejor que otra. Pero si la actividad de ese departamento sirve para otras misiones, es posible que no se pueda distinguir su contribución real.

Otra pregunta que se debe plantear es ¿para qué necesitamos componentes? El ideal para el analista sería que todo fuera una unidad, que existiera sólo el sistema, sin tener que dividirlo en subsistemas. Pero esto es imposible. En consecuencia, la razón real para la separación del sistema en componentes (desde el punto de vista del análisis del sistema) es para proveer al investigador con el tipo de información necesario para

diagnosticar apropiadamente el sistema y decir lo que haya que hacer después.

Desde el punto de vista de las misiones del sistema puede ser útil indicar aquí las diferentes funciones que debe llevar a cabo un sistema para sobrevivir (constituye la misión más importante del sistema total). Katz y Kahn en un interesante estudio de la conducta de las organizaciones desde el punto de vista de la teoría de sistemas, distinguen cinco misiones fundamentales. Ellas son:

1. **La misión de producción.** Es decir, la conversión de la energía en el bien y/o servicio característico del sistema.
2. **La misión de apoyo.** Es decir, las funciones por las cuales se provee de suficiente energía al proceso de producción; la función de "comerciar" la corriente de salida en el medio y así originar las nuevas corrientes de entrada (el ciclo de actividad y las funciones de legalización del sistema en su medio).
3. **La misión de mantención.** Es decir, las funciones destinadas a lograr que los componentes del sistema permanezcan dentro del sistema, cuando éste los requiere, tanto física como psicológicamente.
4. **La misión de adaptación.** Es decir, las funciones destinadas a observar los cambios que se suceden en el medio, predecir las consecuencias que éstos tendrán para el sistema y proponer las medidas necesarias para adaptar el sistema a las nuevas condiciones del medio.'
5. **La misión de dirección.** Es decir, el gobierno del sistema, la coordinación de los subsistemas, la adjudicación de los recursos entre ellos, todo esto para cumplir la misión (o las misiones) general del sistema total.

LA DIRECCIÓN DEL SISTEMA

Si revisamos lo que hasta ahora ha hecho nuestro análisis con el fin de definir con precisión lo que es un sistema total, observaremos que ha definido sus objetivos reales, (es decir, aquellos que determinan la actuación del sistema). Enseguida se enfrentó con el problema de determinar el medio que rodea al sistema, lo que, sin duda, le sirve para conocer la extensión de su unidad de análisis. En tercer lugar examinó los recursos reales y potenciales con que cuenta el sistema para conseguir sus objetivos y, finalmente, como vimos en el punto anterior, procedió a definir las partes, componentes o subsistemas que constituyen el sistema. Ahora ha llegado al último paso, en esta tarea, la administración o dirección del sistema.

Para los efectos del analista de sistemas, esta es aquella parte en donde se generan los planes para el sistema. Es su "inteligencia" y su central de decisiones. Allí es donde se consideran todos los aspectos que hemos discutido en los puntos anteriores. La dirección fija los objetivos de los componentes, distribuye los recursos y controla la actuación y el comportamiento del sistema.

Ahora puede aparecer algo paradojal para el investigador de sistemas. Después del trabajo realizado, puede pensar que si él ha sido el que ha estado trabajando en la construcción del modelo, si han sido sus análisis los que han determinado los objetivos del mismo, el medio, los recursos y los componentes. ¿Acaso entonces no es él el administrador? ¿Debería "amotinarse y hacerse cargo de la dirección con la ayuda de todas sus técnicas y sistemas?".

La verdad es que esto no es así, y él no lo desea tampoco. El es un hombre de ideas, no un hombre de acción. Este último debe correr riesgos y si fracasa no sólo puede ser expulsado de la organización, sino que también puede quedar arruinado. Nuestro investigador de sistemas toma riesgos individuales. Si él fracasa, no tiene que responsabilizarse por el fracaso de todo el sistema.

La administración del sistema no sólo debe generar los planes que éste debe desarrollar, sino también asegurarse de que los planes sean implementados de acuerdo con las ideas originales. Y si no es así, debe determinar el por qué. Esta actividad se conoce generalmente bajo el término de control, aunque la palabra en sí encierre un sentido negativo. Estamos acostumbrados a la idea de control coercitivo, como alguien señalaba, al control del carabinero, que está presto a castigar a quien no cumple. Sin embargo, el concepto moderno de control (especialmente el desarrollado por la ciencia de la información y del control, la cibernetica) lo plantea más bien en el sentido del piloto que controla la dirección del avión. Así, gran parte del control dentro **del** sistema opera a través del principio de excepción, de modo que la administración no interfiere en las actividades de las partes componentes a *menos* que la conducta de éstos evidencie una desviación demasiado grande de los planes.

Sin embargo, el control no sólo significa el examen del correcto desarrollo de los planes y metas señaladas, sino que también implica su evaluación y los consecuentes cambios de planes y metas. Estos cambios constituyen uno de los aspectos más críticos de la administración de sistemas, porque nadie puede decir que se han fijado objetivos correctos, o se ha definido el medio en forma precisa, si se han determinado bien los objetivos y la definición de los componentes. Siempre es probable que se haya escapado algo, que se haya cometido un error, o que se hayan sobrevaluado ciertos recursos. Por lo tanto, la administración del sistema debe recibir informaciones tales que le indiquen cuando su concepto y definición de sistema es errado y debe ser cambiado.

El capitán de un barco, siguiendo la comparación que hace N. Wiener entre esta función de la administración y el gobierno de un barco, tiene la responsabilidad de asegurar que el barco llegará a su puerto de destino dentro de un tiempo prescrito y de acuerdo con un plan de navegación (escalas).

Este es el objetivo general del sistema. Su actuación se medirá en términos de esa meta. El medio del barco es el conjunto de condiciones externas que la nave debe enfrentar: el tiempo, la dirección del viento, la fuerza de las olas, las corrientes marinas, etc.

Desde el punto de vista del capitán, también puede considerarse parte del medio la actuación de las máquinas y la tripulación, ya que éstos se encuentran dados durante el viaje (y siempre que la respuesta a la primera pregunta

- ¿Puedo hacer algo?- sea negativa, y la de la segunda pregunta
- ¿Tiene relación con mi objetivo? - sea afirmativa).

Los recursos del barco son las máquinas, la tripulación, el combustible. Los componentes del sistema son las misiones de la sala de máquina, las misiones de mantención (de máquinas y hombres) de vigilancia, etc. El capitán del barco, como administrador general, da el plan para las operaciones de la nave y vigila su desarrollo correcto. El posee varios sistemas de informaciones dentro del barco que le informan si han ocurrido y dónde han ocurrido desviaciones del plan fijado. Su tarea es determinar el por qué, evaluar el comportamiento de la nave y luego, si es necesario, modificar los planes si la información de que dispone le indica que es aconsejable hacerlo. Esto equivale, como hemos visto anteriormente, a un sistema con circuito cerrado con retroalimentación negativa (generalmente). Un problema central para este tipo de control es la velocidad a que debe fluir la información para que este sistema sea operativo. Cualquiera que haya dirigido una embarcación en un mar "picado" reconocerá que si uno responde en forma demasiado rápida o, al revés, demasiado lenta ante la acción de una ola, se le presentarán problemas. Lo que requiere es una retroalimentación en circuito cerrado que permita al administrador reaccionar a los cambios de las variables externas (o del medio) de una manera óptima.

Notas sobre la teoría general de sistemas

Amparo GARCÍA CUADRADO

Prof. titular de la EUBD. Universidad de Murcia

La gran vigencia que en la actualidad tiene la Teoría de Sistemas como método científico nos lleva a plantearnos cuál ha sido su origen y su desarrollo, así como las dificultades terminológicas que ella conlleva.

A partir de la Teoría General, el universo observado y por tanto los diversos campos científicos no constituyen sectores de saber aislados, sino que todas las ciencias pueden ser consideradas un gran sistema universal de conocimiento donde se dan interdependencias y relaciones. Por ello, la T.G.S. postula, como expondremos más adelante, la necesaria comunicación entre los especialistas de los diferentes campos, de manera que si analizamos, por ejemplo el caso de la Ciencia de la Documentación, disciplina eminentemente interdisciplinaria, observamos como en una de las parcelas de trabajo más interesantes, la referente a la planificación de sistemas de información, es absolutamente imprescindible la puesta en común de conocimientos diversos. «Algunos campos, como la dirección, el comportamiento de las organizaciones, la ingeniería industrial, la informática, la ingeniería eléctrica, las comunicaciones, la psicología y otros, tienen que hacer importantes contribuciones para realizar el estudio y diseño de los sistemas de información»¹.

Así, los sistemas de información, medios a través de los cuales la Documentación intenta alcanzar su objetivo, son inseparables de esta Teoría. Como bien ha expresado Lucas, el análisis y diseño de sistemas de información resulta inseparable de este antecedente conceptual que señala la importancia de examinar cada una de las partes del sistema.

¹ LUCAS, H.C.: *Sistemas de información. Análisis y diseño. Puesta a punto*. Madrid: Paraninfo, 1987, p.18.

Junto a ello, este modelo o propuesta metodológica resulta enormemente fructífera para abordar cada una de las ciencias en sí mismas como sistemas. A modo de ejemplo podemos señalar como el mundo del Derecho es un vasto sistema que abarca, tanto la legislación general de cada rama del derecho: el civil, el penal, el procesal, el administrativo, etc., como el cumplimiento o no de esa legislación, y en consecuencia los órganos jurisdiccionales, y los órganos ejecutivos y de construcción, etc. Por consiguiente el mundo del Derecho es un gran sistema, con muchos subsistemas, tantos como aconseje la misma complejidad de los hechos jurídicos, y de las ciencias que ellos tratan.

Por todo ello parece conveniente examinar esta Teoría General, su origen y desarrollo.

Antes de adentrarnos en el estudio y exposición de esta Teoría conviene establecer qué entendemos por sistema, ya que el sistema constituye el concepto sobre el cual gira esta Teoría General.

En sentido amplio, el *sistema* puede ser concebido como un conjunto de elementos interrelacionados e interactivos. No se trata pues, de un mero conjunto de elementos, que no contengan más que la simple suma de los mismos; conjunto para el cual es válido el principio de que «el todo no contiene más que la suma de sus partes». El sistema es algo más, puesto que a la suma de sus elementos añade:

- a) las mutuas relaciones que los ligan entre sí, y
- b) las acciones y reacciones mutuas de unos elementos sobre otros.

Por eso, para el sistema, vale más bien el principio aristotélico de que «el todo es más que la simple suma de sus partes».

Un sistema es, por ejemplo, el conjunto de los planetas que giran alrededor del Sol juntamente con el Sol mismo. El Sol y los planetas que lo circundan no están simplemente sumados, sino ordenados, con relaciones fijas de distancia de cada uno de los planetas respecto del Sol, dentro de los límites de las distintas órbitas, y además impulsados por movimientos de rotación y traslación, que son también fijos, y de los que resultan otros fenómenos, como son, por referirnos sólo a la tierra, el día y la noche y las cuatro estaciones del año. Por eso sería absurdo o completamente inexacto, el afirmar que el sistema solar se reduce a la suma de la estrella llamada Sol y los planetas denominados Mercurio, Venus, Tierra, Marte, etc. Es verdad que tal suma se da en nuestro sistema planetario, pero no sólo eso, sino bastante más. Retengamos pues, estas dos notas propias de los sistemas, y que los distinguen de los meros conjuntos: la interrelación de sus miembros y la interacción de los mismos.

Para aclarar más el concepto de sistema convendría establecer las diferencias entre el *sistema* y la *estructura*. En los últimos años ha estado muy de moda hablar de estructuras, hasta el punto de dar lugar a una corriente de pensamiento que se conoce con el nombre de *estructuralismo*. También la estructura es un conjunto de elementos enlazados y relacionados entre

sí. Pero no es esencial a la estructura el que sus elementos actúen los unos sobre los otros. Y si bien es cierto que a veces se ha hablado de estructuras dinámicas contraponiéndolas a las estáticas, el papel que fundamentalmente se ha asignado a las estructuras es, más bien el de conformaciones estables, permanentes, o sea, el de esquemas fijos que se mantienen con independencia de los propios elementos, que pueden variar o ser sustituidos por otros. Por ello, las estructuras tienen más bien, carácter estático. En cambio, los sistemas parecen reclamar la interacción, o al menos, la interdeterminación entre sus miembros.

Un sistema, por tanto, consiste en lo siguiente²:

- a) unos elementos que lo componen, que están integrados en el sistema, y que cada uno de ellos posee determinadas propiedades;
- b) una interrelación entre los elementos;
- c) un todo, el sistema, que es distinto de la simple adición o suma de los elementos, y cuyas propiedades son propias del sistema, no de los elementos componentes;
- d) una subordinación de todos los elementos al todo organizado que es el sistema, (e incluso una jerarquización de los elementos entre sí pues hay unos elementos más importantes que otros).

Vista ya la noción de sistema y las notas que lo caracterizan, pasamos a enfrentarnos con la Teoría que nos ocupa, pero lo haremos de una manera sucinta y de modo que resulte fácilmente comprensible. Para ello abordamos en primer lugar los antecedentes de la misma.

ANTECEDENTES DE LA TEORÍA

La idea de considerar las distintas partes que conforman el universo como entidades aisladas en interacción, no es un descubrimiento contemporáneo.

E. Currás señala como ya en el año 2.500 a.C. en el ámbito de la Cultura China se dieron ciertos principios que entrañaban nociones relacionadas con el tema que nos ocupa, como es el caso del principio de que «el todo es mayor que cada una de las partes que lo componen»; como es lógico este principio lleva al estudio de cada una de esas partes que constituyen el todo.

En el mundo griego, la filosofía presocrática, también postulará teorías relacionadas con la cuestión, sin embargo debemos esperar a la época de Platón y Aristóteles para descubrir como estas ideas adquieren carta de naturaleza apareciendo por primera vez el término *sistema*³.

² GONZALEZ NAVARRO, F. La teoría General de Sistemas como matriz disciplinar y como método jurídico. *Persona y Derecho*, 21, 1989, p.114.

³ CURRAS, E.: *La información en sus nuevos aspectos: Ciencias de la Documentación*. Madrid: Paraninfo, 1988, p. 141.

En efecto, Aristóteles ya plantea la máxima de que «el todo es más que la simple suma de sus partes». Para Bertalanffy la frase aristotélica es todavía válida como expresión del problema básico de los sistemas⁴. La filosofía clásica abordará cuestiones relacionadas con la noción de sistema y así nos encontramos con el concepto de *orden*. Este supone un conjunto de cosas o elementos relacionados entre sí, pero con especial referencia a uno primero, dentro del conjunto o sistema, que se relaciona con todos los demás y los demás con él. Es verdad que también los elementos secundarios se relacionan entre sí, pero lo hacen por el hecho de relacionarse todos ellos con el elemento primario con el principio de ese orden. Para los clásicos, pues, todo orden se funda en un principio, y por ello distinguían varias clases de orden, correspondientes a los varios principios ordenadores⁵.

Como se ve, la filosofía griega aporta una serie de consideraciones en torno a los sistemas que serán reconocidas por el propio Bartalanffy al afirmar que nociones relativas a la idea de sistema estaban ya latentes en los pensadores clásicos.

Antecedentes de la Teoría los encontramos también durante la Edad Media. Avanzando en el tiempo, el desarrollo de la filosofía sistemática recibió un auge considerable durante la Edad Moderna con nombres como Kant y Hegel⁶.

Pero centrándonos más concretamente en el desarrollo de la T.G.S. tenemos de situarnos a mediados del siglo XIX. En esta época se produce un movimiento filosófico-científico que trata de poner remedio a la unilateralidad de los métodos de investigación científica. La Ciencia, sin más, se había identificado con la Física, hasta el punto que las ciencias no naturales, a saber, las ciencias culturales, sociales o humanas, o se acercaban en sus métodos a las investigaciones físico-matemáticas, o no eran consideradas como verdaderas ciencias.

A esta tendencia opuso Dilthey una tenaz resistencia. No todas las ciencias podían ser tratadas con los mismos métodos. Los métodos físico-matemáticos habían manifestado su eficacia para el cultivo de las ciencias naturales, llegando a desarrollos y descubrimientos admirables; pero a la par, se habían revelado inadecuados para el cultivo de otro tipo de ciencias que tienen por objeto las obras del hombre. Dilthey las llamó *ciencias del espí-*

⁴ BERTALANFFY, L.von.: Historia y situación de la Teoría General de los Sistemas. En *Tendencias en la Teoría General de los Sistemas*. Selección y prólogo de George J. Klir. 3^a reimp. Madrid: Alianza Universidad, 1987, p.29.

⁵ Por lo demás, esos principios ordenadores no eran más que las cuatro causas, a saber, la eficiente y la final que son causas extrínsecas, y la formal y la material, que son intrínsecas. Y de ello resultaban cuatro órdenes: el etiológico, basado en la causalidad eficiente, el teleológico, basado en la causalidad final, el lógico, basado en la causalidad formal, y el hoológico, basado en la causalidad material.

⁶ Curras, E.: Op. cit. p. 141.

ritu y consideró que constituían la otra parte del *globus intelectualis*, radicalmente distinta a la parte que formaban las *ciencias de la naturaleza*. Esta idea, apuntada ya en su *Introducción a las Ciencias del Espíritu* (publicada en 1883), tendría que haberse desarrollado ampliamente en una segunda parte de esa obra que habría llevado por título «Fundamentación gnoseológica de las ciencias del espíritu». Pero este segundo trabajo no llegó nunca a publicarse. En cambio Dilthey fue desarrollando sus ideas en otros muchos libros, y principalmente en el que lleva por título *Construcción del mundo histórico en las Ciencias del Espíritu* (1905). Este autor describió detalladamente lo que hace de estas ciencias (como la historia, sociología, derecho, economía, etc.) un tipo de saber distinto de las ciencias naturales, y exige, por tanto, para ellas un nuevo método.

En resumen, según Dilthey, las *ciencias naturales* buscan la explicación de los hechos que investigan, pero las *ciencias culturales* buscan la comprensión de los hechos sobre que versan⁷.

Pero más adelante se vió que incluso para ciencias incluidas dentro de las llamadas naturales, como es el caso de la biología, y más aún de la psicología, no resultaban tampoco apropiados los estrictos métodos físico-matemáticos, pues éstos prescinden enteramente de la finalidad, mientras que ni la biología, ni menos la psicología, pueden prescindir del carácter directivo o finalístico de sus objetos. Así, se fue preparando el terreno para la aparición de las teorías organicistas y, finalmente, la Teoría General de los Sistemas.

LUDWIG VON BERTALANFFY

El autor que puso los fundamentos y los desarrollos iniciales de la Teoría General de Sistemas, fue precisamente L. von Bertalanffy (1901-1972). Nacido en Viena, fue profesor de Biología en la Universidad de esta ciudad hasta que, en el año 1949 se traslada a Canadá y posteriormente a Estados Unidos donde gozará de un gran prestigio hasta su muerte.

Los motivos que le llevaron a desarrollar su famosa Teoría fueron, en principio, los mismos que habían conducido a la diversificación de los métodos especializados para cada una de las ramas del saber; pero con una idea nueva o pretendidamente nueva, en todo este asunto, a saber, que se mantienen siempre ciertas analogías, o mejor, paralelismos y correspondencias, entre los problemas que cada ciencia plantea y los métodos adecuados para resolver esos problemas.

Dicho de otra manera: tras defender la diversidad de métodos para la diversidad de saberes, era necesario dar un paso más y descubrir que no to-

⁷ Puede verse una buena exposición del pensamiento de Dilthey sobre el particular en la obra de A. López Moreno *Comprensión e interpretación en las Ciencias del Espíritu*. Murcia, 1990.

do es aquí diversidad, sino también hay algo de semejanza o analogía. Y no sólo eso, sino que en todos los saberes hay problemas parecidos, aunque no iguales, por lo cual las soluciones a esos problemas son también parecidas, pero no iguales. Pues bien, esa semejanza de problemas y esa semejanza de soluciones es lo que se quiere expresar con la T.G.S.

Vamos a ver a continuación algunos aspectos de esta Teoría General siguiendo la exposición del propio Bertalanffy.

a) Planteamiento general

La Ciencia moderna se encuentra hoy caracterizada por la especialización siempre creciente, debida a una inmensa proliferación de datos, la complejidad de las técnicas y de las estructuras teóricas dentro de cada campo. Todo ello trae como resultado que la ciencia se encuentra escindida en numerosas disciplinas que a su vez y de manera constante generan nuevas subdisciplinas.

Ante esta realidad, el hombre de ciencia, el físico, el psicólogo, el sociólogo, etc. están inmersos en sus universos privados y es difícil establecer relaciones entre ellos.

Frente a este aislamiento nos encontramos con que en la ciencia moderna se da un fenómeno curioso: han surgido problemas y concepciones similares en campos muy distintos, de manera independiente⁸.

Para el autor, «este paralelismo de principios cognoscitivos generales en los diferentes campos es aún más impresionante cuando se tiene en cuenta que se dieron independientemente, sin que casi nunca interviniéra nada de la labor e indagación en campos aparte»⁹.

Esto sólo resulta explicable si consideramos que «existen modelos, principios y leyes aplicables a sistemas generalizados o a sus subclases, sin importar su género particular, la naturaleza de sus elementos componentes y las relaciones o fuerzas que imperen entre ellos». Parece pues necesaria la existencia de «una teoría no ya de sistemas de clase más o menos especial, sino de principios universales aplicables a los sistemas en general. De aquí, que adelantemos una nueva disciplina llamada Teoría general de los Sistemas»¹⁰.

b) Dificultades de la Teoría

En un primer momento la T.G.S. fue recibida por el mundo científico con incredibilidad y así, el mismo Bertalanffy nos dice que guardó las notas que sobre dicha teoría había presentado en 1937 en el Seminario filosófi-

⁸ BERTALANFFY, L. von.: *Teoría General de los Sistemas. Fundamentos, desarrollos, aplicaciones*. 6^a reimp. México: Fondo de Cultura Económico, 1987, p. 30.

⁹ Idem., p.31.

¹⁰ Idem., p.32.

co de Charles Morris en la Universidad de Chicago y, no fue hasta después de la Segunda Guerra Mundial, cuando aparecieron sus primeras publicaciones sobre el tema¹¹.

Según nos sigue contando, durante esos años, se habían producido cambios sorprendentes en el clima intelectual y se habían puesto de moda la construcción de modelos y las generalizaciones abstractas. Por otra parte, un nutrido grupo de científicos estaban siguiendo líneas de pensamiento parecidas y habían llegado a conclusiones similares, partiendo de campos diversos y a través de trabajos independientes. Así, Kenneth Boulding escribía lo siguiente a Bertalanffy en 1953: «He llegado casi a sus mismas conclusiones, aunque en mi caso acercándome desde la economía y las ciencias sociales, en lugar de la biología; a saber, que hay un *corpus* de lo que yo he llamado «teoría empírica general», o de «teoría general de sistemas», si utilizamos su excelente terminología, que es aplicable con gran generalidad a un sinnúmero de disciplinas diversas»¹².

Es decir, en palabras del propio autor, «...la Teoría General de Sistemas no estaba tan aislada, ni era una ideosincrasia personal en el grado que yo había creído, sino que correspondía a una tendencia del pensamiento moderno»¹³.

Efectivamente, la Cibernetica, la Teoría de la Información, que introdujo el concepto de información como magnitud medible y desarrolla los principios de su transmisión, la Teoría de los Juegos, la Teoría de la Decisión que analiza decisiones racionales dentro de las organizaciones, las matemáticas relacionales, el análisis factorial y «otros progresos novedosos estaban destinados a enfrentarse a las necesidades de la Teoría General de los Sistemas»¹⁴.

Todo este ambiente intelectual propicio llevó a la constitución de la Sociedad para la investigación de Sistemas Generales, llamada inicialmente Sociedad para la Teoría general de los Sistemas. En 1954, dicha Sociedad estableció un programa de investigación en donde se dice que sus funciones principales son:

«(1) investigar la isomorfía que presenten conceptos, leyes y modelos en varios campos de estudio, y facilitar transferencias útiles entre un campo y otro; (2) impulsar el desarrollo de modelos teóricos adecuados en aquellas esferas donde falten; (3) minimizar la duplicación de esfuerzos en las diferentes disciplinas; (4) promover la unidad de las ciencias mejorando la comunicación entre especialistas»¹⁵. Este programa sigue siendo aún hoy válido como programa de investigación para la T.G.S.

¹¹ Idem., p.94.

¹² Idem., p.13 y también en *Perspectivas en la Teoría General de Sistemas. Estudios científico-filosóficos*. 2º reimp. Madrid: Alianza Universidad, 1986, p.142.

¹³ *Teoría General de los Sistemas*, op. cit., p.93.

¹⁴ Idem., p.93-94.

¹⁵ BARTALANFFY, L. von.: *Perspectivas en la Teoría General de los Sistemas...*, op. cit., p.142-143.

c) Metas

«... la Teoría general de los sistemas es una ciencia general de la «totalidad», concepto tenido hasta hace poco por vago, nebuloso y semimetafísico. En forma elaborada sería una disciplina lógico-matemática, puramente formal en sí misma pero aplicable a las varias ciencias empíricas. Para las ciencias que se ocupan de «todos organizados» tendría una significación análoga a la que disfrutó la teoría de la probabilidad para ciencias que se las ven con «acontecimientos aleatorios»; la probabilidad es también una disciplina matemática formal aplicable a campos de lo más diverso, como la termodinámica, la experimentación biológica y médica, la genética, las estadísticas para seguros de vida, etc.

Esto pone de manifiesto las metas principales de la teoría general de los sistemas:

(1) Hay una tendencia general hacia la integración de las varias ciencias, naturales y sociales.

(2) Tal integración parece girar entorno a una teoría general de los sistemas.

(3) Tal teoría podría ser un recurso importante para buscar una teoría exacta en los campos no físicos de la ciencia.

(4) Al elaborar principios unificadores que corren «verticalmente» por el universo de las ciencias, esta teoría nos acerca a la meta de la unidad de la ciencia.

(5) Esto puede conducir a una integración, que hace mucha falta, en la instrucción científica»¹⁶.

Resumiendo todo lo expuesto podemos extraer las siguientes enseñanzas de esta Teoría general:

a) En primer lugar, tenemos que en todas las ciencias existen sistemas.

b) En segundo lugar, que todos esos sistemas son semejantes o coincidentes en lo fundamental pero no son iguales. El descubrir esas semejanzas y ponderarlas ha sido sin duda el mérito principal de la Teoría.

c) En tercer lugar, tenemos que existen unos principios generales comunes a todos los sistemas. El estudio y consideración de tales principios hará posible la integración de las distintas ciencias y en definitiva un acercamiento a la meta de la unidad de la ciencia.

Para terminar con la exposición de la Teoría que nos ocupa recojamos aquí este juicio realizado por Robert Lilienfeld acerca de las aportaciones de L. von Bertalanffy: «La visión general de von Bertalanffy y sus pretensiones para una significación universal de la teoría de sistemas aparecen ya en sus primeros escritos, que son más bien repetitivos y de índole estática; muchos argumentos se repiten casi textualmente de un libro a otro, lo que

¹⁶ BERTALANFFY, L. von.: *Teoría General de los Sistemas...*, op. cit., p.37-38.

hace al menos que sus concepciones sean consistentes y fáciles de resumir (...) Von Bertalanffy aportó una filosofía programática basada en la biología; se contempló a sí mismo como el iniciador de una nueva e importante senda, aunque delegó en otros el desarrollo y aplicación de sus conceptos en actividades específicas (...)»¹⁷.

APLICACIONES Y DESARROLLOS DE LA T.G.S.

Los planteamientos de Bertalanffy pusieron de actualidad las teorías y los métodos sistémicos. Nombres como Ashby, Shannon, Weiner, Bergman, De Rosnay, Laszlo y otros fueron desarrollando sus diversos campos científicos desde un enfoque sistémico.

Bertalanffy fue consciente de que su propuesta tenía un carácter interdisciplinar y así señala que la T.G.S. es «un modo de ver cosas que antes se habían ignorado o pasado por alto. En este sentido es una máxima metodológica y como toda teoría científica ambiciosa, tiene que ver con los problemas perennes de la filosofía, a los que trata de dar sus propias respuestas»¹⁸.

En efecto, al asignar el término *general*, quiso poner de manifiesto la naturaleza interdisciplinar de la teoría que proponía: todas las ciencias podían ser enfocadas con una perspectiva sistémica.

Pero además, la T.G.S. ha tenido múltiples desarrollos surgiendo la Teoría de Sistemas o Teoría de Sistemas Generales. La diferencia entre ambas es claramente perceptible si tenemos presente que, desde el punto de vista de la T.G.S., en todas las ciencias existen sistemas y estos sistemas son semejantes o coincidentes en lo fundamental pero no son iguales, sino que existen también discrepancias entre los diversos sistemas que es preciso tener en consideración. Por ello, nada tiene de extraño que la T.G.S. haya desembocado, como un complemento necesario de la misma en lo que hoy se denomina Teoría de Sistemas Generales.

Esta T.S. está interesada en categorizar modelos concretos de sistemas, o sea en identificar tipos de sistemas. Por ello el objeto inmediato de esta teoría es el estudio de la totalidad de cada tipo o modelo de sistema, su crecimiento, su diferenciación, su orden jerárquico, su control, sus interrelaciones, su mecánica, su evolución y su finalidad¹⁹.

Los especialistas en esta T.S. han trabajado en la evolución constante de la misma y no han desdenado o no han descuidado ninguno de los aspectos de los sistemas estudiados, y que pueden ayudar a conocerlos y caracterizarlos lo más adecuadamente posible. Por eso dentro de esta T.S. hay que consi-

¹⁷ BARTALANFFY, L. von.: *Perspectivas en la Teoría General de Sistemas*..., op. cit., p.49.

¹⁸ Idem., p. 154-155.

¹⁹ CURRAS, E.: *La información en sus nuevos...* op.cit., p.154.

derar como partes de la misma: una filosofía de sistemas, para estudiar los aspectos filosóficos que puedan descubrirse en ellos; una matemática de sistemas para expresar de modo conveniente las medidas y las correspondencias que en ellos se encuentran; una tecnología de sistemas, para considerar sus aspectos técnicos y su posible construcción; una dinámica de sistemas, un análisis y unos procesos de síntesis; la dialéctica sistémica, la invéntica, etc.

Tal como recoge Bertalanffy en uno de sus escritos, «la ciencia clásica en sus diversas disciplinas, la química, la biología, la psicología o las ciencias sociales, había intentado aislar las partes que componían los universos observados —compuestos químicos, enzimas, células, sensaciones elementales, individuos en libre competencia, etc.—, con la esperanza de que, al reconstruir conceptual o experimentalmente el todo o sistema —célula, mente, sociedad—, éste resultase inteligible. Hoy sabemos que para comprenderlos no sólo son necesarias sus partes, sino también las relaciones que conectan unas con otras y con el entorno: el interjuego de enzimas en la célula, los procesos conscientes o inconscientes de la personalidad, la estructura y la dinámica de los sistemas sociales, etc.»²⁰.

Si como vemos, en un principio se abordó el estudio de los sistemas como conjuntos de elementos interrelacionados entre sí y con el medio, había sin embargo, como señala E. Currás, ciertos comportamientos que no tenían explicación, por lo que hubo que pensar en la existencia de unos procesos dinámicos que regulaban el comportamiento de los sistemas. Así es como hace su aparición la *dinámica sistémica*, parte de la T.S. de enorme aplicación en el campo de la economía, de la política, de los estudios de mercado, etc. En definitiva, la *dinámica sistémica* se interesa por las transformaciones de los sistemas en el tiempo. Para el estudio de la evolución de los sistemas en el tiempo, los analistas utilizan la llamada *simulación sistémica* o construcción de modelos en los que se hacen constar parámetros y variables²¹.

Al aplicar esta dinámica sistémica a un tipo de sistema caracterizado por su evolución discontinua o por etapas de duración variable, se comprobó que no todos los comportamientos sistémicos tenían explicación. Esto ha motivado la aparición de la llamada *dialéctica sistémica* que es definida por Rodríguez Delgado como un método experimental deductivo-inductivo para estudiar las leyes o principios generales que gobiernan los procesos de transformación de ciertos tipos de sistemas²².

Hoy día también se lleva a cabo el estudio de los sistemas por medio de los procesos de análisis y síntesis. El *análisis de sistemas* permite abordar la estructura, el funcionamiento y las relaciones del sistema por medio de la *evaluación sistémica* o metodología en donde la información sobre todos los as-

²⁰ *Perspectivas...* op. cit. p.144.

²¹ CURRAS, E.: Op.cit., p.143 y 157.

²² Idem., p.162.

pectos del sistema resulta imprescindible para poder realizar la evaluación del mismo. Así, a la hora de abordar el desarrollo de un sistema de información es preciso llevar a cabo el análisis del mismo para su posterior diseño o proceso planificador. Es decir, es preciso comprender el sistema en su totalidad llevando a cabo un proceso de clasificación e interpretación de hechos, diagnóstico de problemas, etc, especificando lo que el sistema debe hacer. Junto a ello, los *procesos de síntesis* «tratan de estudiar el sistema en su conjunto, considerando las partes integradas de forma global en el todo»²³. En definitiva, los procesos de síntesis tienen como fin la mejor concepción y organización de los sistemas por medio de la *simulación sistemática* y la *invéntica*.

Por último, se ha desarrollado la *sistemografía* o metodología consistente en descubrir las características del sistema y sus interrelaciones mutuas y con el entorno, las etapas de su evolución, etc. Para Emilia Currás «es este el estado de evolución de la teoría de sistemas en el momento presente con una prospectiva de futuro hacia un desarrollo y auge aún mayor»²⁴.

Pues bien, todos estos estudios juntos y combinados integran una completa y omnicomprensiva Teoría de Sistemas.

CARACTERÍSTICAS DE LA TEORÍA DE SISTEMAS

Machlup y Mansfield han señalado las siguientes notas acerca de la Teoría de Sistemas²⁵, a las que hemos añadido alguna apostilla.

1. «No hay consenso general en cuanto al ámbito de la T.S.» Esto es así, debido en gran parte a la falta de una terminología común entre los varios autores incluso cuando se trata de definir el concepto de sistema. Sin embargo, para F. Machlup «Esto no debe preocuparnos: los eruditos pueden no estar de acuerdo en cómo definir un mismo término y, sin embargo, están de acuerdo en su significación. Son demasiado orgullosos para dar la excusa del profano: «usted sabe lo que quiero decir», cuando es consciente de no poder expresarse con claridad. Y los científicos suelen poder transmitir significados claros sin formular definiciones explícitas; (sin embargo) esto no es así cuando hablan de sistemas. Más aún, tres o cuatro nociones son mencionadas como fundamentales por la mayoría de los teóricos de sistemas: un todo (o reunión, colección, grupo, etc), partes (o elementos, objetos, entidades constructivas, miembros, etc.) relaciones (entre las partes), y a veces, como una idea adicional, el medio...»²⁶

²³ Idem., p.157.

²⁴ Idem., p.143.

²⁵ Notas sintetizadas por A. Debons en una serie de puntos en su obra conjunta con E. Horne y S. Cronenweth, *Information Science. An integrated view*. Boston: G.K. Hall and Co., 1988.

²⁶ MACHLUUP, F. y MANSFIELD, U. (eds): *The study of information Interdisciplinary messages*. New York, etc.: John Wiley & Sons, 1983, p.43.

A este respecto G.J. Klir señala que «una faceta peculiar de la Teoría general de sistemas, es su *terminología*. La terminología de sistemas, aunque aspire a ser el lenguaje propio para la comunicación interdisciplinaria, se reduce en la actualidad a una mezcla poco trabada de lenguajes utilizados por distintos individuos o grupos. Por ejemplo, es de lamentar que al tiempo que hay con frecuencia varios nombres distintos para un mismo concepto, conceptos distintos tengan a veces el mismo nombre. Tales ambigüedades son causa de numerosas confusiones. Además, este caos de términos levanta dudas acerca de toda la teoría general de sistemas» por lo que es necesario llevar a cabo con absoluta prioridad la unificación terminológica²⁷.

En parecidos términos se expresa González Navarro para quien «resulta sorprendente que la T.G.S. que proclama su vocación de universalidad, no se haya preocupado gran cosa, hasta ahora, de precisar su terminología y depurar las unidades conceptuales. Y desde luego, por lo que a mí respecta puedo afirmar que la mayor dificultad que he encontrado para entender la literatura especializada que aquí manejo deriva, no tanto del uso frecuente de fórmulas matemáticas —ciencia ajena a mi dedicación profesional— cuanto de ese hecho de falta de acuerdo previo acerca de los términos empleados y de los conceptos que tratan de designar. El mismo significante se emplea por unos autores con un significado y con significado distinto por otros»²⁸.

2. «La Teoría de Sistemas es primordialmente matemática.» Esta característica no es indiscutible pues hay fenómenos que no pueden ser matematizados ya que la cualidad no se puede matematizar. Sin embargo, el objetivo de la Teoría es el desarrollo de ésta «en términos matemáticos puesto que las matemáticas son un lenguaje exacto que permite deducir, confirmar y rechazar teorías de manera rigurosa»²⁹.

3. «La Teoría de Sistemas no es una teoría de sistemas concretos sino una teoría de modelos.» De aquí que también se la denomine Teoría de Sistemas Generales, sin descender al nivel específico o concreto sino manteniéndose en un nivel genérico con rasgos suficientemente precisos e identificables.

4. «La Teoría de Sistemas contribuye a las otras ciencias generando un conjunto de modelos para cualquier sistema bien definido pero no especifica qué modelo es el mejor para un sistema particular.»

5. «La teoría de sistemas es reduccionista.» En realidad todas las ciencias lo son, pues en la medida en que buscan cierto grado de universalidad

²⁷ KLIR, G. J.: Teoría polifónica general de sistemas. En *Tendencias en la Teoría ...*, op. cit., pp.23-24.

²⁸ GONZALEZ NAVARRO, F.: La Teoría General de Sistemas como matriz disciplinar y como método jurídico. *Persona y Derecho*, 21, 1989, p.109.

²⁹ BERTALANFFY, L. von.: *Perspectivas en la Teoría...*, op. cit., p.45.

o generalidad tienen que prescindir de los datos o de los hechos demasiado individualizados.

6. «La Teoría de Sistemas suministra las bases para organizar las ciencias sociales». Pero no sólo las ciencias sociales, sino todas ellas dado su carácter de matriz disciplinar.

7. «Mientras la Teoría General de Sistemas trata de lo fundamental, abstracto y aspectos generales de los sistemas, la Teoría de Sistemas trata de las cuestiones más específicas acerca de las clases de sistemas mejor definidos».

CLASIFICACIÓN Y DESCRIPCIÓN DE LOS PRINCIPALES MODELOS DE SISTEMAS

Se han establecido diversas clasificaciones acerca de los sistemas teniendo en consideración aspectos diversos como propiedades, funciones, configuración, grado de complejidad, ámbito, etc.

Teniendo en cuenta la simplicidad o complejidad de los elementos que se integran en el todo podemos señalar la existencia de sistemas **simples** y sistemas **complejos**. Estos últimos están integrados por subsistemas, que a su vez pueden tener otros subsistemas y así hasta la unidad básica del sistema que constituye lo que denominamos sistema simple.

Un ejemplo claro lo podemos observar si analizamos el Sistema Archivístico Español. Este sistema complejo, está integrado por dos subsistemas, también complejos, el sistema estatal y el sistema autonómico. El primero de ellos está integrado por dos nuevos subsistemas que son la Red de Archivos de titularidad estatal, adscritos o no al Ministerio de Cultura y los Centros y Servicios archivísticos de carácter técnico. Cada uno de ellos está a su vez compuesto por diversos subsistemas como son, por una parte los archivos Nacionales, archivos Regionales y Provinciales y archivos de la Administración Central (subsistemas de la Red de Archivos) y por otra, el Servicio Nacional de Microfilm, el Instituto de Conservación y Restauración y el Centro de Información Documental de archivos (subsistemas de Centros y Servicios). Cada uno de estos últimos, son unidades básicas del sistema. Del mismo modo cada uno de los subsistemas de tercer orden que son los archivos Nacionales, archivos Regionales y Provinciales y los de la Administración Central, están constituidos por unidades básicas de ese sistema que es el Sistema Español de Archivos.

De igual manera, el otro gran subsistema de primer orden que es el Sistema Autonómico está integrado por 17 subsistemas de las respectivas Comunidades Autónomas que, a su vez, se subdividen en nuevos subsistemas hasta llegar a las unidades básicas. Así, el Sistema Regional de archivos de la Comunidad de Murcia está configurado por dos subsistemas, los archivos públicos y los privados. Los primeros están constituidos por las si-

gientes unidades básicas: el Archivo de la Administración Regional, el Archivo Histórico Provincial y los diversos archivos municipales. En cuanto a los privados tenemos aquellos que, siendo de interés público, reciben ayudas o subvenciones por parte de la Comunidad Autónoma; por tanto cada uno de los archivos privados que reunan estas condiciones constituirían las unidades básicas de ese sistema murciano de archivos.³⁰

Considerando la naturaleza de los sistemas, tenemos que éstos se dividen en sistemas **reales** y sistemas **lógicos**, siendo los primeros aquellos en los que existe verdadera actividad, y los segundos aquellos en los que tal actividad falta por completo.

Los sistemas reales son los que se dan en las cosas mismas y que cada ciencia estudia y los lógicos son aquellos que las distintas ciencias tratan de reflejar, o trasponer en las mentes humanas en relación a los sistemas reales. Estos sistemas lógicos y reales son denominados por Acroff y Emery³¹ como sistemas **abstractos** y sistemas **concretos**. Es decir, aquellos en que todos los elementos que los componen son conceptos son sistemas abstractos; y las entidades específicas que existen en el espacio y en el tiempo son sistemas concretos. Así, el ser humano es un sistema concreto. El sistema de las cifras y las letras por los que nosotros representamos la abstracción matemática y los conceptos son también sistemas concretos; sin embargo la abstracción matemática es un sistema abstracto.

Dentro de los sistemas reales o concretos nos encontramos en primer lugar, con los sistemas **naturales**, o que la propia naturaleza ha construído. Son los sistemas del universo físico, de los astros, de los seres inertes de la tierra, de los seres vivos del mundo vegetal, y de los seres vivos del mundo animal. Por tanto, dentro de estos sistemas naturales podemos distinguir entre sistemas **físicos o inorgánicos** y sistemas **vivos**, llamados también por algunos autores sistemas **orgánicos**. Todos estos sistemas son muy variados pero con ciertas analogías.

Dentro de los sistemas reales tenemos también los que se deben a la intervención del hombre, agente inteligente y libre, en la misma naturaleza, y que constituyen la amplísima y variadísima gama de los seres artificiales o sistemas **artificiales**. Finalmente están los sistemas que resultan de la convivencia de los propios hombres y que podemos llamar sistemas **sociales**; la agrupación en familias, en pueblos, en ciudades, en naciones, en sociedades de todo tipo, juntamente con las vicisitudes históricas de todos estos grupos sociales.

³⁰ Ponencia de Alvarez-Coca, M.J. y Gómez Llera, E. Redes y Sistemas de archivos. Legislación. Organos, Centros. Recursos. *Actas de la IV Conferencia de la ANABAD, La Coruña, 1988. Boletín de ANABAD*, 38, 1-2, 1988, p.9-80.

En cuanto al sistema autonómico de la Región de Murcia: *Murcia. Leyes 1990 Enero-Junio*. Murcia: Consejería de Cultura, Educación y Turismo; Comunidad Autónoma, 1990.

³¹ Según recogen Debons, Hornc y Cronenweth en su obra *Information Science...*, op. cit.

Entre todos estos sistemas hay una analogía. La analogía o semejanza de los sistemas lógicos con los sistemas reales es indudable, puesto que cada sistema lógico trata de trasponer en nuestra mente el sistema real al que se refiere. Analogía, pues, de cada sistema lógico con el sistema real al que trata de reflejar o trasponer. Pero solamente analogía, no igualdad, porque las diferencias entre lo real y lo lógico son grandes e irreductibles; en la realidad hay actividad, pero en lo lógico no la hay. Por eso, un sistema real es un conjunto de elementos que se hallan entre sí, no sólo relacionados, sino también en interacción; unos elementos actúan sobre los otros y viceversa. En cambio, los sistemas lógicos son un conjunto de elementos relacionados entre sí como lo están las conclusiones con las premisas, o sea por relaciones de implicación o de inferencia, sin el menor rastro de verdadera actividad.

Pero hay también analogías entre todos los sistemas reales. Todos son conjuntos de elementos interrelacionados e interactivos, y en todos ellos hay un orden, un orden que descansa en el papel preponderante de uno de tales elementos, el que hace por así decirlo, de centro, o mejor, de principio de todos los demás. Así, en nuestro sistema planetario, el elemento central o principal es el sol; en el sistema circulatorio, el corazón, en el sistema digestivo, el estómago, etc.

Dependiendo de sus relaciones con el medio que les rodea, tenemos: sistemas **cerrados** y sistemas **abiertos**. El sistema abierto es aquel sistema real (concreto) cuyos límites son permeables a la transmisión de materia, energía o datos del entorno y desde el sistema al entorno, es decir, establece relaciones de intercambio con el medio que le rodea. Es esta una característica general de los seres vivos, sistemas inexorablemente abiertos. Este planteamiento es defendido por Bertalanffy en su *teoría de sistemas abiertos*. Dice así el autor: «Todo organismo viviente es un sistema abierto, que se caracteriza por importar y exportar substancias sin descanso. En este intercambio el organismo rompe y reconstruye sus elementos, pero se mantiene constante. Es lo que yo he llamado estado estable...»³². Es decir, cualquier sistema viviente está abierto a la materia (como el comer) y a la energía (a través de los sentidos) y de retorno crea sonidos, emite calor, elimina residuos, etc.³³. O lo que es lo mismo, recibe entradas y produce salidas.

Pero debemos considerar también como sistemas abiertos, otros sistemas que no siendo sistemas vivos, es decir, orgánicos, son capaces de producir entradas y salidas y mantener ese estado estable. Es el caso por ejemplo, de un horno regulado por un termostato que le permite mantener un calor estable. Este horno es evidentemente un sistema abierto, que mantiene relaciones con su entorno de donde toma energía; sin embargo no es

³² *Perspectivas en la teoría general de sistemas...*, op. cit., p.40.

³³ *Information Science: An integrated view...*, op. cit.

un sistema vivo³⁴. Hemos de concluir pues, que no todos los sistemas abiertos son sistemas vivos, mientras que todo sistema orgánico o viviente es un sistema abierto.

Otro ejemplo de este tipo de sistema abierto es el sistema de información. Este sistema mantiene relaciones con su entorno, y recibe de él información. Esta información actúa en dos sentidos: por una parte se genera en el interior información, por medio de la cual, el sistema se comunica con el exterior, con otros sistemas, con el ambiente que le rodea. Pero además, la información alimenta el sistema desde fuera, es decir, desde el medio que le rodea. Por tanto se establece un flujo-reflujo de información, que en ningún caso es reversible. La información que fluye hacia el sistema de información no es la misma que la que refluye o emana del sistema. Las entradas continuas de datos y documentos, evita que el sistema se destruya y las demandas de información desde el exterior, deben estar en consonancia con las respuestas para mantener el equilibrio del sistema. En efecto, un sistema de información «que no guarde un equilibrio entre los fondos recibidos y su actividad informativa, terminará destruyéndose por asfixia o por convertirse en obsoleto»³⁵.

El sistema cerrado es aquel que no intercambia materia con el medio ambiente. Hablando en términos estrictos, no se puede admitir la existencia de tales sistemas. El único sistema cerrado, sería el universo en su conjunto, ya que cualquier parte del universo se encuentra relacionada con las otras partes, y muchas veces con intercambio activo de energías de todo tipo. Para Senn, el sistema cerrado no existe en la realidad, es un simple concepto³⁶.

Se podrían seguir enumerando otros tipos de sistemas ya que como señala Currás el mundo de la Teoría de Sistemas es tan extenso como imaginar pueda la mente humana. Para concluir con este punto vamos a encuadrar dentro de la tipología expuesta las características de un sistema de información.

Todo sistema de información es un sistema real, es decir, una entidad específica que existe en el espacio y en el tiempo; es un sistema artificial, pues su existencia se debe a la acción del hombre con objeto de que el fin de la Documentación pueda desarrollarse. Al mismo tiempo es un sistema abierto ya que mantiene relaciones con su medio ambiente, como ya hemos analizado (científicos, técnicos y demás productores de información, usuarios del sistema, legislación, tecnología, economía, etc.). Por otra parte el sistema de información puede ser simple, cuando los elementos que lo integran

³⁴ GONZALEZ NAVARRO, F.: La Teoría General de Sistemas como matriz disciplinar... op. cit., p.116

³⁵ CURRAS, E.: Op. cit., p. 152.

³⁶ SENN, J. A.: *Análisis y diseño de sistemas de información*. 2º ed. México, etc.: Mc Graw-Hill, Cop., 1992, p. 21.

son individualidades, o complejos, cuando sus elementos no son individualidades básicas, sino otros sistemas que, con respecto al sistema más amplio en el que se integran, pueden y deben denominarse subsistemas³⁷.

CONCLUSIÓN

Para terminar con esta exposición hemos de señalar que la Teoría General de Sistemas está en continua evolución y desarrollo, siendo numerosos los especialistas que han expuesto sus enfoques particulares sobre la Teoría. Ello ha originado concepciones diversas de la misma. Para unos es una teoría formal, para otros una metodología, una forma de pensar y analizar el universo, e incluso una herramienta educativa y una profesión³⁸. A pesar de la diversidad de puntos de vista hay algo común a todos los autores que por ella se han interesado y es el reconocerle su carácter innovador. Es innovadora porque varias razones:

1. Permite observar el mundo como un conjunto de fenómenos individuales pero interrelacionados en lugar de aislados;
2. Demuestra que ciertos conceptos, principios y modelos no dependen de la naturaleza específica de los fenómenos implicados;
3. Abre a través de investigaciones generales, nuevas posibilidades a disciplinas específicas. Es importante que a través de la T.G.S. se reconozca el valor específico de otras ciencias, distintas de las naturales o fisicomatemáticas, que puedan exhibir la seriedad, el rigor y la exactitud que hoy se reclama para todos los demás;
4. Resalta la unidad del saber humano y fomenta así los estudios interdisciplinarios, que tanto se echan a veces de menos.

³⁷ Para más información sobre taxonomía sistemática consultar el libro de E. Currás *La información en sus nuevos aspectos...* op.cit., págs. 147-153.

³⁸ Sobre esta última cuestión George J. Klir insiste en la necesaria formación de generalistas especializados y de especialistas generalizados. «Con frecuencia se llama generalistas a quienes trabajan en la teoría general de sistemas, contraponiéndoles a los *especialistas* que trabajan en alguna disciplina clásica. Sin embargo, quien trabaja únicamente en la teoría general de sistemas, se convierte en un especialista. Se especializa en generalizaciones. Llamémosle *teórico de sistemas o generalista especializado* (...) Afirmo que la existencia de un gran teórico de sistemas capaz de resolver casi todo problema de casi toda disciplina, es un mito, y creo que como tal debe tratarse. Un teórico de sistemas se especializa en investigación de los principios generales de los sistemas, y unas pocas horas, días o incluso semanas de estudio concentrado de otra disciplina, no pueden darle sino una comprensión muy superficial de sus peculiaridades, necesidades y problemas (...) Un teórico de sistemas no puede dominar todas las materias en las que va a trabajar hasta el punto de resolver cualquier problema especializado que pueda surgir. Pero un especialista en, por ejemplo, sanidad, puede dominar los fundamentos de la teoría general de sistemas en un tiempo relativamente pequeño. Diremos que se trata de un *especialista generalizado*.» Teoría polifónica general de sistemas. En *Tendencias en la teoría...*, op. cit., p.23.

Capítulo 2

Teoría de Control

Capítulo 2

Teoría de Control¹

Los objetivos de este capítulo son:

- Conocer todos los conceptos relacionados a la Teoría de Control.
- Aplicar estos conceptos en ejemplos simples.

2.1 Introducción

Tal como se dijo en el capítulo anterior, el universo es un gran sistema formado por muchos subsistemas los cuales le suplen de metas y por ende de autocontroles. La creación del universo, como ese indescriptible sistema dotado de un gran número de otros sistemas (formando una jerarquía de sistemas), nos da la idea de la necesidad de un gran y único control en ellos, desde el principio, para que permanezcan hasta la actualidad.

El concepto de control ha jugado un papel vital en el desarrollo de la ciencia y de la ingeniería, además de su integración e importancia en los procesos industriales y de manufacturas modernas, así como de gerencia y administración del hogar. Su teoría y práctica brindan medios de lograr el funcionamiento óptimo de sistemas dinámicos, mejorar la calidad, abaratar los costos de producción, expandir el ritmo de producción; en fin, mejorar la eficiencia de los sistemas.

La tendencia actual en la ingeniería de sistemas es hacia una mayor complejidad, debido principalmente a los requerimientos de tareas complejas, y de una buena precisión. Debido a la necesidad de afrontar los cada vez más severos requisitos del comportamiento del sistema de control, al aumento en su complejidad, y al fácil acceso a computadoras de gran escala se introduce entonces el concepto de *teoría de control*, para poder estudiar más fácilmente estos sistemas tan complejos.

2.2 ¿Qué es Teoría de Control?

La teoría de control se basa en el principio de gobernar un proceso, esto es, dirigir una variable determinada a niveles o necesidades establecidas mediante acciones que se deciden, en función de las señales emitidas por dichas variables, sobre los elementos del proceso que han de gobernarse o controlarse. A veces lo que se controla es un sistema, en otras palabras, el proceso principal (si hay más de uno) del sistema.

¹ Material revisado gentilmente por el Ing. Lino Ruiz

En esta teoría, para que exista el control, debe haber:

- Un objeto, ente o factor.
- Un subsistema controlador.
- Algo que se necesite controlar en el objeto (variable por controlar en el objeto),
- Un observador o detector que se encargue de revisar las señales que se reciban de la variable controlada,
- Un componente comparador que examine la discrepancia entre el valor actual de la variable controlada y lo deseado.
- Una reacción a lo comparado.

En todo sistema de control deben estar presentes estos seis elementos. Un ejemplo que explicaría la función de estos elementos en un sistema sería el siguiente:

Un atleta, después de haber recorrido una distancia, empieza a experimentar un cambio en su temperatura corporal y en ese instante comienza a excretar sudor por los poros. Los elementos de este ejemplo son:

- **Un objeto:** El atleta después de haber recorrido una distancia.
- **Subsistema Controlador:** La excreción por los poros.
- **Variable por controlar:** La temperatura corporal.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Puede ser un termómetro o algún aparato que sirva para observar la variación de la temperatura, a través del tiempo en que el atleta está excretando o sudando.
- **Comparador:** Desviación = temperatura corporal (29° C) versus temperatura deseada (27.5° C).
- **Reacción:** Sudar para bajar la temperatura².

La teoría de control también es conocida como Ingeniería de Control. Esta, en lugar de operar con dispositivos y componentes físicos, se le reemplaza por sus modelos matemáticos. Uno de los problemas más importantes en Ingeniería de Control es obtener un modelo matemático razonablemente exacto de un sistema. Teoría de control es cualquier interconexión de componentes que satisfacen una función deseada. También es conocida como sistema de control retroalimentado, que tiende a mantener una relación preestablecida entre las salidas y algunas entradas de referencia, comparándolas y utilizando la diferencia como medio de control. Los sistemas de control no están limitados al campo de la ingeniería, sino que se le puede encontrar en todas las disciplinas del conocimiento. Por ejemplo, el organismo humano es análogo a un sistema de control muy avanzado³.

También se le llama servosistema o servomecanismo en el que la salida es algún

² Puede ser temblar en caso que la temperatura corporal sea menor que la temperatura deseada; esto quiere decir que hace frío.

³ En si se controlan muchos factores

elemento mecánico, sea posición, velocidad o aceleración.

Algunas de las ventajas que ofrece la teoría de control son:

- Incremento en la exactitud. El sistema se puede diseñar para tratar de llevar a cero el error entre la respuesta deseada y medida.
- Pequeña sensibilidad a los cambios en los componentes. El sistema puede diseñarse para tratar de obtener cero errores, a pesar de los cambios en los componentes.
- Reducir los efectos de las perturbaciones. Se pueden atenuar notablemente los efectos de perturbaciones al sistema.
- Incremento en la rapidez de respuesta y anchura de banda. La teoría de control puede utilizarse para incrementar la gama de frecuencias sobre la cual un sistema responderá y hacer que responda más satisfactoriamente. Por ejemplo, un cohete impulsor de satélite tiene todas aquellas semejanzas aerodinámicas de un gigantesco palo de escoba.

La teoría de control es ampliamente difundida y utilizada en todas partes y cada vez juega un papel más importante en el desarrollo de las sociedades y de las tecnologías. Es utilizada en cualquier área científica, industrial, comercial, administrativa, tecnológica y hasta social.

El ejemplo del ser humano, que hemos citado anteriormente, es quizás uno de los sistemas de control más sofisticado y complejo que existe. Un ser humano promedio es capaz de llevar a cabo una gran diversidad de tareas, incluyendo la toma de decisiones. Algunas de estas tareas, como la recolección de objetos o una caminata, suelen ser labores rutinarias. Bajo ciertas circunstancias, estas tareas deben ejecutarse en forma óptima. Por ejemplo, un atleta que corre 100 metros planos tiene por objetivo recorrer esta distancia en el menor tiempo posible. Un corredor de maratón no sólo debe correr la distancia con la mayor rapidez, sino que, además, para lograrlo, debe controlar el consumo de energía y obtener un resultado óptimo. Por consiguiente, se puede decir en forma general que la vida impone el logro de "muchos objetivos", y los medios para alcanzarlos casi siempre dependen de sistemas de control.

Los sistemas de control ejercen mucha influencia sobre cada faceta de la vida moderna. Las secadoras, satélites espaciales, plantas de proceso químico, sistemas de orientación y navegación, control de la contaminación, tránsito colectivo y regulación económica son unos cuantos ejemplos. En el sentido más amplio, un sistema de control puede ser definido como una interconexión de varios componentes, todos trabajando juntos para la consecución de un objetivo común. Estos componentes, en su mayoría, son dispositivos eléctricos, mecánicos, electrónicos y electromecánicos, entre otros. En otros casos son variables de algún problema abstracto.

En la industria, por ejemplo, son comunes los sistemas de control. Desde el control de calidad de productos industriales, líneas de ensamble automático, tecnología espacial y armamento, control por computadora, sistemas de transportación, robótica y muchos

otros. También se pueden resolver con enfoques de teoría de control automático, problemas como el control de inventarios y los sistemas de control sociales y económicos.

Cualquiera que sea el tipo de sistema de control bajo consideración, los elementos básicos del sistema pueden describirse gráficamente a través de diagramas de bloques, que en forma elemental.

En el caso de un sistema de control los elementos más significativos de su composición son:

- Objetivos del control
- Componentes del Sistema de Control
- Proceso o planta del sistema
- Resultados

Esto se ilustra en el siguiente diagrama de bloques:

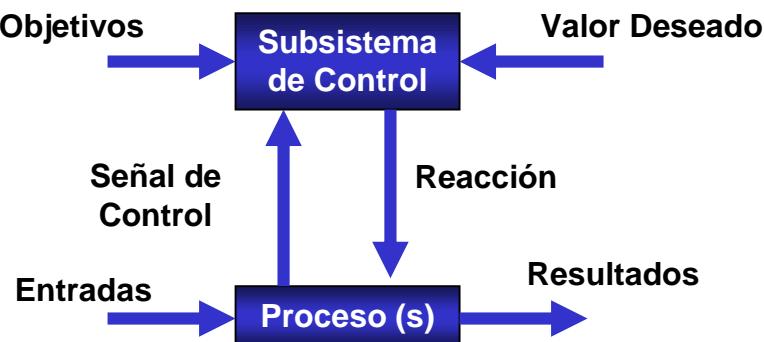


Figura 2.1: Diagrama de Bloque de un Sistema de Control

En general, el objetivo de un sistema de control se basa en controlar las salidas de una manera predeterminada, por medio de las entradas y aplicando los elementos del sistema de control. El sistema de control posee unas entradas⁴ y resultados⁵. Las entradas al subsistema de control se denominan *señales de control* y las salidas como *reacciones*. Existen sistemas en los que los controles están acoplados. A los sistemas con más de una entrada y una salida se les llaman *sistemas complejos de control*.

Control implica medir el valor de la variable controlada del sistema, y aplicar al sistema la variable manipulada para corregir o limitar la desviación del valor medido, respecto al valor deseado. La variable controlada es la cantidad o condición que se mide y controla; esta es normalmente la salida del sistema. Con el fin de afectar la variable controlada, la

⁴ estímulos

⁵ respuestas

cantidad o condición modificada por el controlador, se conoce como variable manipulada.

2.3 Principios Básicos de Diseño de Sistemas de Control

En esta sección se presentan algunos puntos que son relevantes cuando se diseña un sistema de control.

2.3.1 Requerimientos Generales de un Sistema de Control

Los requerimientos generales de un sistema de control son los siguientes:

1. Debe ser capaz de reducir errores a cero, o a un valor pequeño tolerable.
2. La velocidad de respuesta debe ser razonablemente rápida.
3. Todo sistema de control debe ser relativamente estable, es decir, la respuesta debe mostrar un amortiguamiento razonable.

Cualquier sistema de control, para ser útil, debe satisfacer estos requisitos. El requisito de estabilidad relativa razonable y el de la precisión tienden a ser incompatibles. Por lo tanto, al diseñar sistemas de control resulta necesario efectuar el mejor compromiso entre estos dos requerimientos. Un ejemplo que explicaría estos requerimientos podría ser el funcionamiento de un edificio inteligente:

El desarrollo e implementación de la inteligencia de un edificio tiene como objetivo primordial el confort y la seguridad de las personas que habitan o trabajan en el mismo. Protegiendo física y psicológicamente el entorno humano circundante, utilizando todos los recursos posibles para el cometido final, analizando el límite y la reducción posible de los costos de explotación.

Ejemplos de controles (facilidades) de un edificio inteligente:

- Control y Optimización de Energía Eléctrica
- Sistema de Alimentación Interrumpida
- Control de Iluminación
- Control de Elevadores
- Sistema de Protección Contra Incendios
- Sistema de Control de Aire Acondicionado
- Sistema Integrado de Seguridad
- Sistema de Gestión y Control de Habitaciones
- Sistema de Control de Estacionamientos
- Sistemas de Comunicaciones
- Sistemas de Distribución de TV
- Megafonía
- Busca Personas
- Sistema de Control de Salas de Reuniones y Salas Multifuncionales
- Cableado Estructurado del Edificio

2.3.2 Teoría de Control Clásica versus Teoría de Control Moderna

Teoría de Control Clásica

Se basa en la relación entrada-salida del sistema, o función de transferencia. Sólo se consideran importantes las señales de entrada, salida y error; se desarrolla el análisis y diseño de sistemas de control utilizando funciones de transferencia juntamente con una variedad de técnicas gráficas.

Este método se aplica únicamente a sistemas lineales invariantes en el tiempo, de una entrada y una salida. Las técnicas utilizadas son conceptualmente simples y sólo requieren una cantidad razonable de cálculos. Lamentablemente, el proyecto de sistemas en la teoría de control clásica se basa en procedimientos de tanteo y ajuste, los que en general no producen sistemas de control óptimos.

Teoría de Control Moderna

Se basa en el concepto de estado, el cual existe desde hace mucho tiempo en el campo de la dinámica clásica y en otros campos. El estado de un sistema dinámico es el conjunto más pequeño de variables tales que el conocimiento de esas variables en $t = t_0$, juntamente con la entrada para $t \geq t_0$, determinan totalmente el comportamiento del sistema para cualquier tiempo $t \geq t_0$.

En otras palabras, el comportamiento futuro de un sistema, quedará determinado por su historia pasada y por las excitaciones externas que le afecten.

Este método es aplicable a sistemas de múltiples entradas y múltiples salidas que pueden ser lineales o no lineales, invariantes o variables en el tiempo.

El diseño de sistemas, por este método, permite al ingeniero proyectar sistemas de control óptimos con respecto a los índices de comportamiento dados. Además, el proyecto se puede realizar para toda clase de entradas en lugar de que se realice para una función específica de entrada. La Teoría de Control Moderna se utiliza para analizar los sistemas de control, de forma más viable, general y eficaz que la Teoría de Control Clásica.

La teoría de control moderna es un nuevo procedimiento para el análisis y diseño de sistemas de control complejos, la cual está basada en el concepto de estado y espacio de estado. Los sistemas complejos pueden tener múltiples entradas y múltiples salidas y ser variables en el tiempo.

La teoría de control clásica brinda, generalmente, buenos resultados para sistemas de control de una entrada y una salida. Los procedimientos clásicos o convencionales, ponen énfasis en la comprensión física y utilizan menos cálculos matemáticos que los métodos de control modernos.

La teoría de control moderna contrasta con la teoría de control clásica, en que la primera se aplica a sistemas de múltiples entradas y múltiples salidas, que pueden ser lineales o no lineales, variables o invariantes en el tiempo, mientras que la segunda tiene mayor aplicación en sistemas lineales, invariantes en el tiempo, y de una sola entrada y una sola salida. Además, la teoría de control moderna es un procedimiento en el dominio del tiempo esencialmente, mientras que la teoría convencional opera en el dominio de las frecuencias complejas, y por tal razón utiliza ampliamente técnicas de solución numéricas de ecuaciones diferenciales.

2.3.3 Modelado Matemático

Para poder programar en la computadora un sistema de control, se requiere plantear el problema en función de un modelo matemático. En ingeniería de control, se usan ecuaciones diferenciales lineales, invariables en el tiempo, funciones de transferencia y ecuaciones de estado, para modelos matemáticos de sistemas lineales, invariables en el tiempo y de tiempo continuo.

Para ser útil, un modelo matemático, este no debe ser ni muy complicado, ni excesivamente simple. Un modelo matemático debe representar los aspectos esenciales de un componente físico. El uso de tales modelos matemáticos permite a los ingenieros de control desarrollar una teoría de control unificada.

2.3.4 Análisis y Diseño de Sistemas de Control

Para poder diseñar un sistema de control es necesario ejecutar una metodología que posea todos los pasos necesarios para lograr un sistema eficiente. Por ejemplo, el control automático resulta esencial en el control numérico de las máquinas y herramientas en las industrias manufactureras. También resulta esencial en operaciones industriales como el control de presión, temperatura, humedad, viscosidad, y flujo en las industrias. La Ingeniería Mecánica Industrial diseña y construye sistemas de este tipo.



Figura 2.2: Ensamblaje de Automóviles

Fuente: www.bjinforma.com

Como los avances en la teoría y práctica del control automático brindan medios para lograr el funcionamiento óptimo de sistemas dinámicos, mejorar la productividad, liberarse de la monotonía de muchas operaciones manuales rutinarias y repetitivas, y otras ventajas; la mayoría de los ingenieros y científicos deben poseer un buen conocimiento de este campo.

2.3.5 Método Básico de Diseño de Control

El método básico para diseño de cualquier tipo de sistemas prácticos de control, necesita la aplicación obligada de procedimientos de tanteo.

La especificación de la señal de control durante el intervalo de tiempo operativo recibe el nombre de ley de control. Matemáticamente, el problema básico de control es determinar la ley de control óptimo, sujeta a diversas restricciones de ingeniería y de economía, que minimice un índice de comportamiento o desempeño determinado. Para sistemas de control aeroespacial, el índice de comportamiento puede ser un mínimo en tiempo o en combustible. Por ejemplo, para sistemas de control industrial, el índice de comportamiento puede ser el costo mínimo, la confiabilidad máxima, tiempos muertos mínimos, etc. Es importante puntualizar que la elección del índice de comportamiento es sumamente importante, ya que la naturaleza del control óptimo diseñado depende del índice de comportamiento particular que se elige. Así pues, se debe seleccionar el índice de comportamiento más adecuado para cada situación.

2.3.6 Pasos de Diseño

Tras completar el diseño matemático, el ingeniero de control simula el modelo en una computadora, para verificar el comportamiento del sistema resultante en respuesta a diversas señales y perturbaciones. Generalmente, la configuración del sistema inicial no resulta satisfactoria. Luego, se debe rediseñar el sistema, y completar el análisis correspondiente. Este proceso de revisión de diseño y análisis se repite hasta obtener un sistema satisfactorio. Al cabo, ya se puede construir un prototipo del sistema físico. Este tema se trata en mayor detalle en el siguiente capítulo.

2.4 Clasificación de los Sistemas de Control

La distinción de los sistemas de control la determina la acción de control, que es la cantidad que activa el sistema para producir la salida. Los sistemas de control de maneta más general se pueden clasificar en:

- Sistemas de control de **ciclo abierto**.
- Sistemas de control de **ciclo cerrado**.

2.4.1 Sistema de Control de Ciclo Abierto

Los *sistemas de control de ciclo⁶ abierto* son aquellos sistemas en los cuales la acción

⁶ también conocido como *lazo*.

de control es independiente a la salida, o sea, la salida no se mide ni se retroalimenta para compararse con la entrada.

El sistema de control que se muestra en la figura 2.2, es poco sofisticado y se clasifica como *sistema de control de ciclo abierto*. Resulta fácil apreciar que dicho sistema no cumplirá en forma satisfactoria con los requerimientos de desempeño deseados.

Debido a la simplicidad y economía de los sistemas de control de ciclo abierto, se usan en la práctica en muchas situaciones. De hecho, casi todos los autos fabricados antes de 1981 no contaban con un sistema de control.

Los elementos de un sistema de control de ciclo abierto, casi siempre pueden dividirse en dos partes: el controlador y el proceso controlado, tal como se muestra en la figura 2.3, se aplica una señal de entrada o comando al controlador, cuya salida actúa como señal de control u ; la señal actuante controla el proceso controlado, de tal manera que la variable controlada c se comporte de acuerdo con estándares predeterminados. Estos estándares de comportamiento, en la práctica se determinan experimentalmente y luego se programan las acciones de control, para producir estos efectos.

Ejemplos:

- El control de tráfico por señales actuadas en función de tiempo.
- Una lavadora de las llamadas “automática” es un sistema de ciclo abierto, puesto que el tiempo de lavado se ha determinado mediante criterios y cálculos previos. Una lavadora, verdaderamente automática, debería comprobar constantemente el grado de limpieza de la ropa y desconectarse por sí misma cuando dicho grado coincidiese con el deseado.
- Un motor de gasolina que se utiliza para accionar una gran bomba, como se ilustra en la figura 2.4, el carburador y la máquina comprenden un tipo común de sistemas de control en donde una gran potencia de salida se controla mediante una pequeña potencia de entrada. En este caso el carburador es el controlador y el motor es el proceso controlado. La proporción de combustible es el control de entrada, y la bomba de carga es una señal perturbadora.



Figura 2.3: Ciclo Abierto

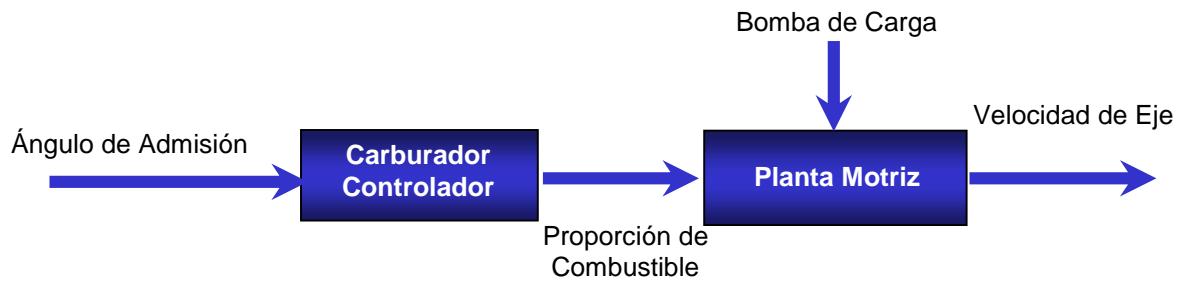


Figura 2.4: Motor de Gasolina

2.4.2 Sistemas de Control de Ciclo Cerrado

Es aquel sistema en el cual la acción de control es en cierto modo dependiente a la salida. Estos sistemas son también llamados sistemas de control por *retroalimentación*⁷.

En los sistemas de control de ciclo abierto, el elemento faltante para lograr un control más preciso y adaptable es un ciclo cerrado o retroalimentación de la salida a la entrada del sistema. Para obtener un control más preciso, la señal controlada $c(t)$ debe retroalimentarse y compararse con la entrada de referencia, tras lo cual se envía a través del sistema una señal de control proporcional a la diferencia entre la entrada y la salida, con el objeto de corregir el error o la desviación. A los sistemas con uno o más ciclos de retroalimentación de este tipo se les llama *sistema de ciclo cerrado*.

Sin embargo, resulta paradójico en este tipo de sistema, que para que exista una acción de control, primero debe ocurrir un error del sistema. Esto implica, que se pueda lograr mejorar la precisión, pero deberá considerarse que la acción de control sea oportuna y no desfasada.

En la figura 2.5 se muestra el diagrama de bloques de un sistema de control de ciclo cerrado aplicado al control de velocidad de un motor. La entrada de referencia w_r fija la velocidad deseada. Comúnmente, cuando el par de carga es cero, la velocidad del motor en reposo debe coincidir con el valor de referencia w_r , y cualquier diferencia entre la velocidad real y el valor deseado, causada por cualquiera perturbación del par de carga tl , es detectada por el transductor de velocidad y detector de errores, con lo que el controlador operará sobre esta diferencia y proporcionará una señal para ajustar el ángulo del obturador a que corrige el error.

Ejemplos:

- Máquinas frigoríficas.
- Calderas para la generación de vapor.

⁷ La palabra *retroalimentación* también se conoce como *realimentación*.

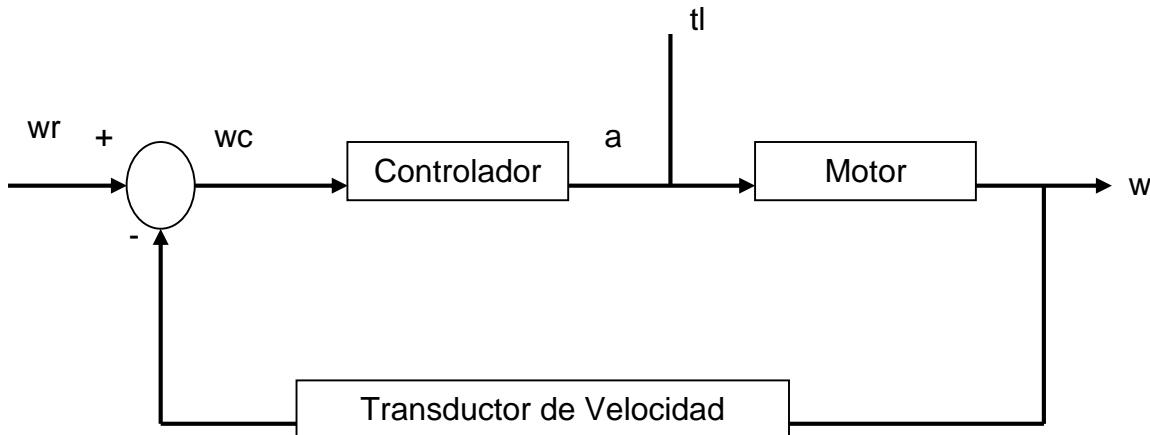


Figura 2.5: Sistema de Control de Ciclo Cerrado

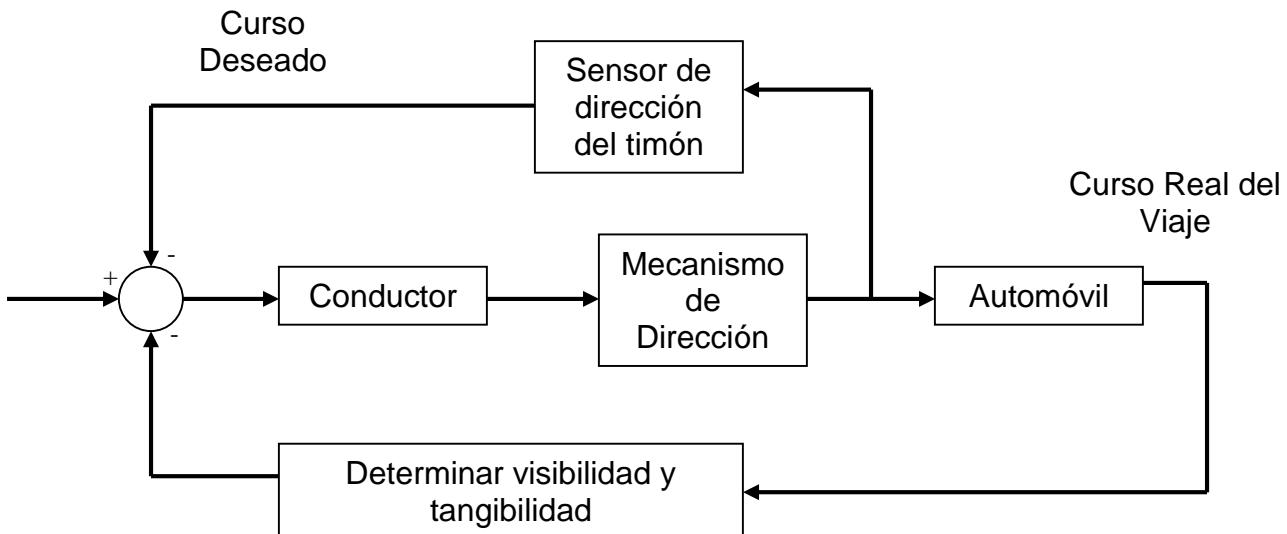


Figura 2.6: Sistema de Dirección de un Automóvil

- Sistema de control de conducción de un automóvil. En la figura 2.6 se presenta el diagrama de bloques de un sistema de control para la conducción de un automóvil. El curso deseado es comparado con la medida de error. Esta medida se obtiene por la retroalimentación visual y táctica (movimiento corporal). Este sistema de retroalimentación es una versión similar al sistema de manejo en una línea marítima o el control del vuelo en un gran avión. Todos estos sistemas operan en una secuencia de ciclo cerrado. También pueden ser conducidos con el apoyo de satélites y sistemas de posicionamiento global.

2.5 Tipos Básicos de Sistemas de Control de Ciclo Cerrado

Los sistemas de control de ciclo cerrado son los más utilizados en la actualidad y para facilitar su estudio presentamos los principales tipos que se han desarrollado.

2.5.1 Control Directo

Es recomendable medir y controlar directamente la variable que indica el estado del sistema o calidad del sistema para obtener un mejor resultado. Para el caso de control de proceso se puede desear, medir y controlar directamente la calidad del producto. Esto puede presentar un problema, ya que puede ser difícil efectuar la medición de calidad. Por ejemplo, se pueden controlar variables como, temperatura y presión, directamente relacionadas con la calidad.

2.5.2 Control Indirecto

El control indirecto de un sistema suele no ser tan eficaz como el control directo, ya que puede haber otras variables que afecten la relación entre la calidad y las variables medidas. Esto quiere decir que, por medio de terceros, se controla el sistema.

2.5.3 Sistema de Control Adaptable

Las características dinámicas de la mayoría de los sistemas de control no son constantes por diversas razones, como el deterioro de los componentes al transcurrir el tiempo o las modificaciones en parámetros o en el medio ambiente. Aunque en un sistema de control retroalimentado se atenúan los efectos de pequeños cambios en las características dinámicas, un sistema para ser satisfactorio ha de tener la capacidad de adaptación, lo cual implica la capacidad de auto ajustarse o auto modificarse de acuerdo con modificaciones imprevisibles del medio o estructura. Los sistemas de control que tienen algún grado de capacidad de adaptación, se denominan *sistemas de control adaptados o adaptables*.

En un sistema de control adaptado, las características dinámicas deben estar identificadas en todo momento, de manera que los parámetros de control o detección puedan ajustarse para mantener el funcionamiento óptimo. Este concepto tiene una buena dosis de atracción para el ingeniero de sistemas, ya que un sistema de control adaptable, además de acomodarse a los cambios ambientales, también lo hace ante moderados errores de proyecto de ingeniería o incertidumbre y compensa la eventual falla de componentes menores del sistema, aumentando la confiabilidad de todo el sistema.

2.5.4 Sistemas de Control con Aprendizaje

Si intentamos analizar sistemas de control con operación humana, encontramos el difícil problema de escribir ecuaciones que describen el comportamiento humano. En este caso

uno de los muchos factores que lo complican es la capacidad de aprender del operador humano. A medida que este va adquiriendo experiencia, se convierte en un elemento de control, y esto debe ser tomado en cuenta al analizar el sistema. Los sistemas de control con capacidad de aprender reciben el nombre de *sistemas con aprendizaje*. En los últimos años se han dado varios avances en lo referente a estos sistemas.

2.5.5 Sistemas de Regulación Automática

Es un sistema de control retroalimentado en el que la entrada de referencia o salida deseada pueden ser constantes, o bien, variar lentamente en el tiempo, y donde su tarea fundamental consiste en mantener la salida en el valor deseado a pesar de las perturbaciones presentes. Algunos ejemplos de este sistema son: el regulador centrífugo de Watt, la regulación automática de tensión en una planta generadora ante variaciones de carga eléctrica, y los controles automáticos de presión y temperatura en un proceso químico.

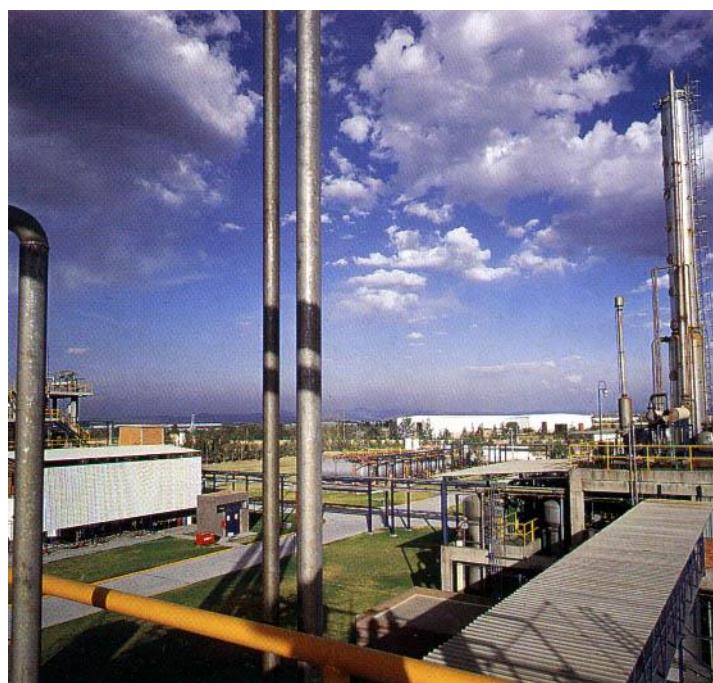


Figura 2.7: Planta Química

Fuente: www.canacintra.org.mx

2.5.6 Sistemas de Control de Procesos

Es un sistema de regulación automática en el que la salida es una variable como temperatura, presión, flujo o nivel de líquido. Este sistema tiene amplia aplicación en la industria. Con frecuencia se utilizan en estos sistemas controles programados, como el de la temperatura de un horno de calentamiento en el que la temperatura del mismo se controla según un programa preestablecido.

Por ejemplo, el programa preestablecido consiste en elevar la temperatura a determinado valor durante un intervalo de tiempo, luego reducir a otra temperatura prefijada también durante un período predeterminado. Así pues, nos damos cuenta que el controlador funciona manteniendo la temperatura del horno cercana al punto de ajuste variable.

2.6 Comparación entre los Sistemas de Control

En este punto analizaremos los pros y contras de los diferentes sistemas de control que existen en nuestro tiempo.

2.6.1 Ciclo Abierto versus Ciclo Cerrado

Una ventaja del sistema de control de ciclo cerrado es que el uso de la retroalimentación hace al sistema, en su respuesta, relativamente insensible a perturbaciones externas y a variaciones internas de parámetros del sistema. De este modo es posible utilizar componentes relativamente inexactos y económicos y poder lograr la exactitud de control requerida en determinada planta; mientras esto sería imposible en el caso de ciclo abierto.

Desde el punto de vista de la estabilidad, en el sistema de control de ciclo abierto es más fácil de lograr, ya que la estabilidad no constituye un problema importante. Por otro lado, en los sistemas de ciclo cerrado la estabilidad siempre constituye un problema de importancia, por la tendencia a sobre corregir errores, que puede producir oscilaciones de amplitud constante o variable. Hay que recalcar que para sistemas en los que las entradas son conocidas previamente y en los que no hay perturbaciones, es preferible usar el control de ciclo abierto. Los sistemas de ciclo cerrado solamente tienen ventajas si se presentan perturbaciones no previsibles y/o variaciones imprevisibles de componentes del sistema.

Generalmente, se logra un funcionamiento satisfactorio y más económico de todo el sistema, si se opta por una combinación adecuada de controles de ciclo abierto y cerrado.

2.6.2 Lineales versus No Lineales

Estos sistemas se basan en los métodos de análisis y diseño. En su concepto más estricto, los sistemas lineales no existen en la práctica, pues todos ellos tienen un cierto grado de no-linealidad. Los sistemas de control lineales retroalimentados son modelos idealizados que sólo existen como concepto en la mente del analista, para simplificar el análisis y diseño. Cuando las magnitudes de las señales de un sistema de control están limitadas en un intervalo en el que los componentes exhiben características lineales, el sistema es esencialmente lineal.

Por ejemplo, un efecto no lineal de los sistemas de control es la asimetría o desajuste mecánico de los miembros acoplados con engranajes, las características no lineales de los resortes, las fuerzas de fricción o torsión no lineales entre miembros móviles, entre otras. Con frecuencia las características no lineales se introducen a propósito en los

sistemas de control para mejorar su desempeño o lograr un control más efectivo. Por ejemplo, para obtener un control de tiempo mínimo se usa un controlador de tipo cierre-apertura. Este tipo de control es frecuente en muchos sistemas de control de proyectiles o vehículos espaciales.



Figura 2.8 Transbordador

Fuente: www.latercera.cl

En el caso de los sistemas lineales, existe una gran diversidad de técnicas analíticas y gráficas para propósitos de diseño y análisis. Sin embargo, el tratamiento matemático de los sistemas no lineales es bastante difícil y no se cuenta con métodos generales que puedan aplicarse a la resolución de un grupo amplio de sistemas no lineales.

2.6.3 Invariable en el Tiempo versus Variable en el Tiempo

Un sistema de control invariante en el tiempo es aquel en el que los parámetros no varían en el tiempo. La respuesta de tal sistema es independiente del tiempo en el que se aplica la entrada.

Un sistema de control variable en el tiempo es aquel en el cual los parámetros varían con el tiempo; su respuesta depende del tiempo en el que se aplica una entrada. Un ejemplo de un sistema de control variable en el tiempo es el *control de un proyectil* dirigido durante un vuelo. Aunque un sistema variable con el tiempo sin linealidad es todavía un sistema lineal, el análisis y diseño de esta clase de sistemas suelen ser mucho más complejos que los sistemas lineales invariantes en el tiempo.

2.6.4 Con Parámetros Concentrados versus con Parámetros Distribuidos

Los sistemas de control que pueden describirse mediante ecuaciones diferenciales ordinarias, son llamados *sistemas de control con parámetros concentrados*, mientras que los sistemas de control que se describen mediante ecuaciones diferenciales parciales se llaman *sistemas de control con parámetros distribuidos*.

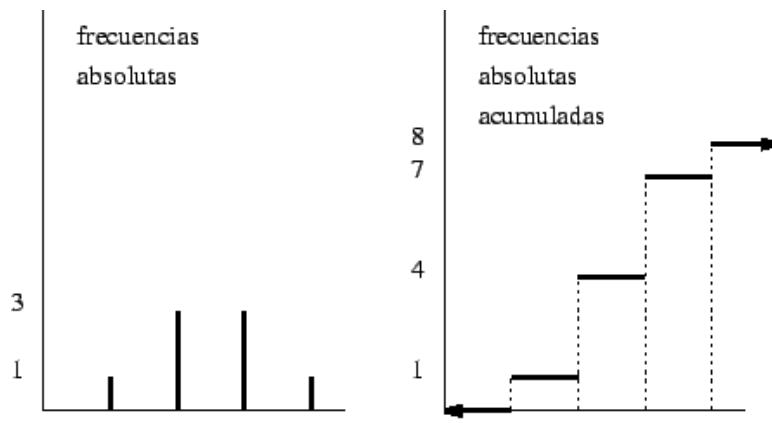
2.6.5 Determinísticos versus Estocásticos

Un sistema de control es determinístico si la respuesta a la entrada es predecible y repetible. De no ser así, si el sistema involucra probabilidades (incertidumbre), es estocástico.

2.6.6 Tiempo Continuo versus Tiempo Discreto

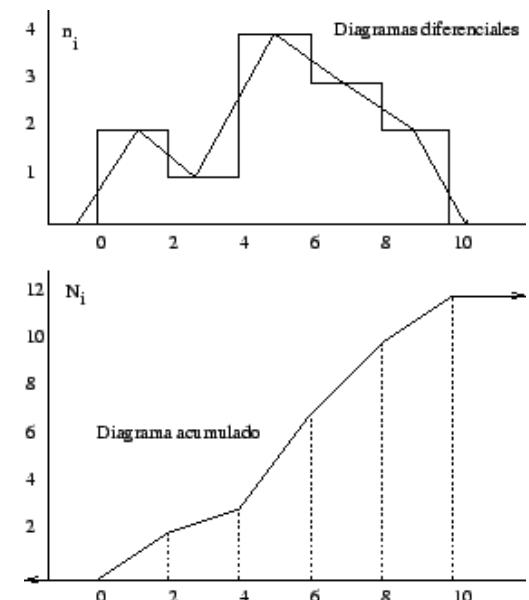
En un sistema de control de tiempo continuo, todas las variables son funciones de un tiempo continuo t . Un sistema de control de tiempo discreto, abarca una o más variables que son conocidas sólo en instantes discretos o paquetes de tiempo.

A continuación, ejemplos de gráficas representativas de este tópico:



Diagramas diferenciales e integrales para una variable discreta.

Fuente: www.bioestadistica.uma.es



Diagramas diferenciales e integrales para una variable continua.

Fuente: www.bioestadistica.uma.es

2.6.7 Con una Entrada y una Salida versus con Múltiples Entradas y Múltiples Salidas

Un sistema puede tener una entrada y una salida. Por ejemplo, un sistema de control de posición, donde hay un comando de entrada (la posición deseada) y una salida controlada (la posición de salida). Algunos sistemas pueden tener múltiples entradas y múltiples salidas. Otro ejemplo, un sistema de control de proceso en 2 entradas: Entrada de presión y entrada de temperatura y 2 salidas: Presión de salida y temperatura de salida.

2.7 Problemas de Control

De una manera informal, el problema de control consiste en seleccionar, de un conjunto específico o arbitrario de elementos (o parámetros, configuraciones, funciones del tiempo, etc.), aquellos que, aplicados a un sistema fijo, hagan que este se comporte de una cierta manera deseada.

Los problemas de control muestran tres elementos en común:

- Uno que se puede modificar, llamado *entrada*.
- Uno que se desea que tenga ciertas características, llamado *salida*.
- Un tercero, llamado *meta*, que relaciona la entrada con la salida y que no puede ser modificado.

Podemos decir, que el problema de control consiste en escoger, para un sistema dado, una entrada que haga responder a la meta de una manera deseada, esto es, que sostenga una salida con cierta característica. En la siguiente tabla podemos resumir tres problemas de sistemas de control con sus respectivos elementos:

Problema	Entrada	Salida	Meta
Control de Nivel	Localización del punto de operación	Variaciones en el nivel del líquido	Relaciones mecánicas del sistema
Inversionista	Cantidad de acciones a comprar y vender en cierta fecha	Cantidad de efectivo al finalizar el año	Mecanismo de la Bolsa de Valores
Reactor Químico	Temperatura del flujo de alimentación	Composición del torrente de salida	Relaciones de balance y cinética del reactor

2.8 Ejemplos

A continuación, se presentan una serie de ejemplos que utilizan los principios y técnicas de la teoría de control.

2.8.1 Control de Velocidad del Automóvil

El conductor es el encargado de determinar cuál será la velocidad apropiada para la situación. Un generador eléctrico produce un voltaje proporcional a la velocidad dada. Este voltaje se compara con un voltaje de referencia (el correspondiente a la velocidad anterior) que corresponda a la velocidad deseada. La diferencia en los voltajes es utilizada como señal de error para alterar el carburador e incrementar o decrementar la velocidad según se requiera.

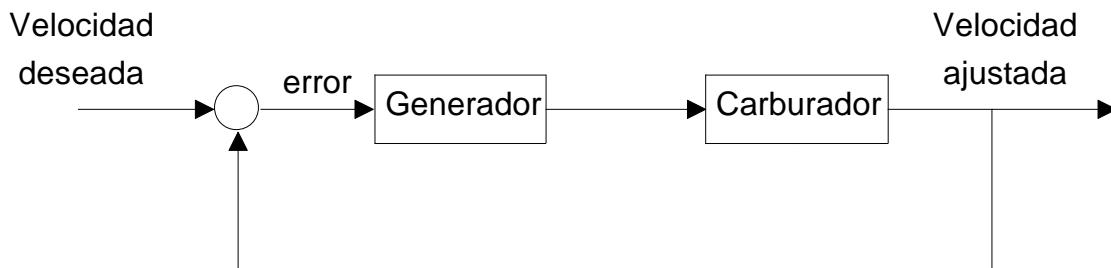


Figura 2.9: Control de Velocidad de un Automóvil

Componentes:

- **Objeto:** El automóvil después de haber recorrido una distancia.
- **Subsistema Controlador:** Conductor automático⁸.
- **Variable por controlar:** La velocidad del automóvil.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Velocímetro que observa la variación de la velocidad, a través del tiempo en que el conductor está acelerando.
- **Comparador:** Desviación = velocidad actual (90 km/h) versus velocidad deseada (80 km/h).
- **Reacción:** Aumentar o disminuir la velocidad.

2.8.2 Sistema de Control para llenar un Recipiente con Agua

Objetivo:

Por medio de un grifo en la base, el agua se vaciará en el recipiente. El sistema cortará el flujo de agua automáticamente cuando el recipiente se llene.

⁸ Conocido en inglés como *automatic cruiser*.

Componentes:

- **Objeto:** El recipiente después de haber abierto el grifo.
- **Subsistema Controlador:** El sistema formado por el tapón y el flotador.
- **Variable por controlar:** El nivel del agua.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** El tapón el cual nos sirve para observar, a través del tiempo el nivel del agua en el recipiente.
- **Comparador:** Desviación = nivel actual (10 m) versus nivel máximo (10 m).
- **Reacción:** El tapón cerrará el flujo de agua de la tubería base.

Funcionamiento:

Tendremos una entrada de agua desde una tubería base, hacia el recipiente que se quiere llenar. En el recipiente tendremos un flotador que ascenderá a medida que el recipiente se esté llenando.

Atado al flotador, existe una cuerda en cuyo extremo opuesto tiene un tapón que controla la entrada de agua desde la tubería base. Cuando el flotador llega al límite de llenado del recipiente, el tapón cerrará el flujo de agua de la tubería base (Véase la figura 2.10).

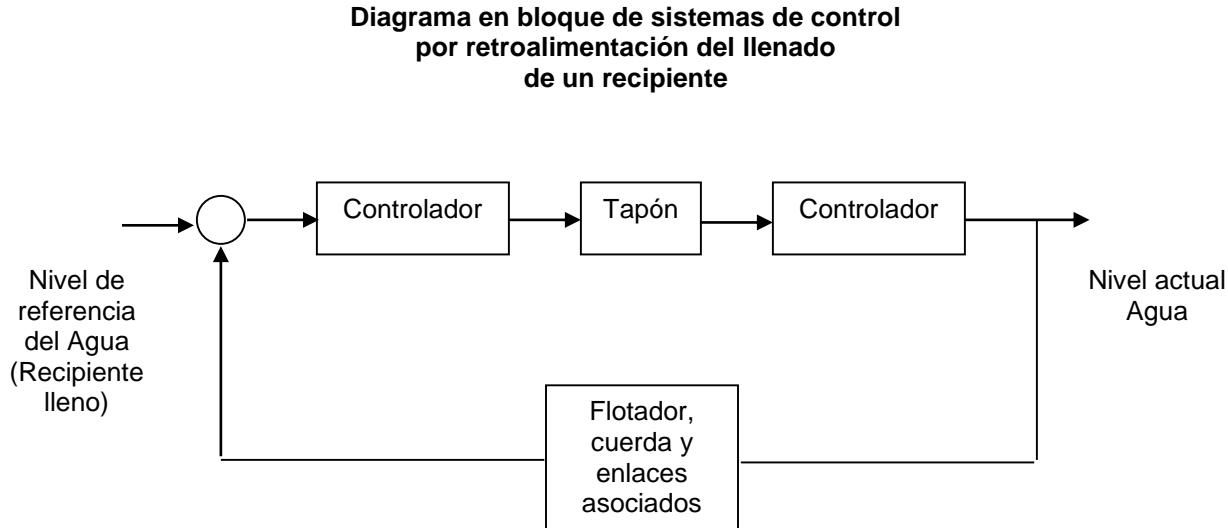


Figura 2.10: Recipiente de Agua

2.8.3 Sistema de Control Biológico formado por un Ser Humano al tomar un Objeto.

Los componentes básicos de este sistema de control son el cerebro, el brazo, la mano y los ojos.

El cerebro envía la señal del sistema nervioso requerida hacia el brazo y la mano, con el

fin de alcanzar el objeto. Esta señal se amplifica en los músculos del brazo y la mano los cuales sirven como impulsores en el sistema. Los ojos se usan como dispositivos de exploración y continuamente “retroalimentan” hacia el cerebro la información sobre la posición de la mano. La salida del sistema es la posición de la mano. La entrada es la posición del objeto.

El objetivo del sistema de control es reducir la distancia entre la posición de la mano y la del objeto a cero. La figura 2.11 es un diagrama esquemático. La línea punteada y las flechas representan la dirección del flujo de información.

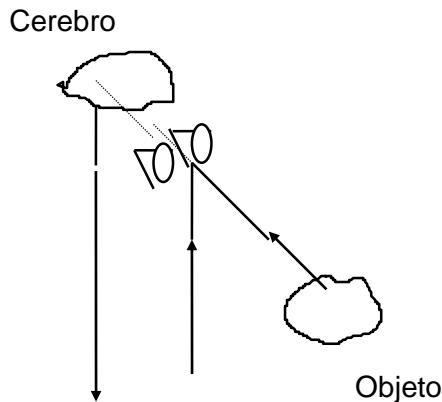


Figura 2.11: Ser Humano que toma un objeto

Componentes:

- **Objeto:** Persona que desea tomar un objeto.
- **Subsistema Controlador:** El cerebro.
- **Variable por controlar:** La distancia entre la mano y el objeto que se desea agarrar.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Los ojos los cuales nos sirven para observar, a través del tiempo, la distancia a la que nos encontramos del objeto que se desea agarrar.
- **Comparador:** Desviación = distancia actual (8 m) versus distancia deseada (10 m).
- **Reacción:** Acercarse más al objeto, ya sea extendiendo más los brazos o caminando más cerca del objeto.

2.8.4 Sistema de Información

Auditoría en un Contexto Computarizado

Objetivo:

Integrar los conceptos de factores y riesgos de control en un sistema computarizado y su incidencia en el enfoque de auditoría que se deseé aplicar.

Enunciado:

La siguiente es una descripción resumida de la aplicación computarizada de cobranzas de cuentas corrientes de Computer Company al 31 de marzo de 1994.

La aplicación de cobranzas de cuentas corrientes fue adquirida del estudio de consultoría MSC y Asociados, que vendió el producto llave en mano, es decir, en funcionamiento y sin que los usuarios ni los miembros del centro de cómputo tengan acceso a los programas.

La cajera, al cabo de cada día, ingresa al sistema la totalidad de los recibos que le entregan los cobradores. Previamente los totaliza en la forma manual y arma un comprobante con los recibos del día.

Para ingresar al sistema debe colocar su identificación de usuario y su clave de acceso que le fue asignada en enero de 1995, fecha de la implantación del sistema. La cajera debe ingresar el total del comprobante antes de comenzar a cargar cada recibo en particular. A medida que esto se produce, el sistema controla la información de los recibos validando, entre otros, los siguientes datos: número de cliente, número de factura e importe de la factura.

Además, el sistema se inhabilita para el procesamiento de dos cobranzas por la misma factura, ya que la modalidad de cobro consiste en recaudar la totalidad de las facturas. Cuando las cobranzas son rechazadas por el sistema, la cajera las coloca en una carpeta para el análisis diario. Cuando el total del comprobante cargado al sistema balancea con la suma de los recibos ingresados, tiene lugar la actualización de los archivos de cuentas corrientes. Los comprobantes diarios de recibos son enviados para su proceso en lote al Centro de Informática mediante el ingreso al sistema de tesorería y de contabilidad, afectando, entre otras, las cuentas de caja, bancos, cuentas locales y del exterior.

Mensualmente el supervisor de contaduría concilia el listado de cuentas con la cuenta control mayor, investigando las diferencias y dejando constancia de su aprobación.

Solución:

Puntos débiles de control	Riesgo afectado
La clave de acceso no se cambia habitualmente	Acceso a la aplicación

Puntos fuertes de control	Riesgo disminuido
Controles de validación en el ingreso de los recibos	Ingreso de datos
Controles de totales sobre los ingresos de los recibos	Ingreso de datos
El sistema inhabilita el doble procesamiento de las facturas	Ingreso de datos

Enfoque de la Auditoria Propuesta:

- Revisión de la conciliación mensual del auxiliar de cuentas versus el mayor.
- Revisión de los comprobantes rechazados al cierre de un período y del adecuado seguimiento por parte de la cajera.

Componentes:

- **Objeto:** Sistema de información computarizado
- **Subsistema Controlador:** Sistema de Auditoría (ejemplo: ACL)
- **Variable por controlar:** Un campo de una tabla de la base de datos
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Programa computacional que indague el campo de la tabla.
- **Comparador:** Programa computacional que compare el valor del campo para ver si cumple con sus características (ejemplo: campo tipo moneda)
- **Reacción:** Reporte de Inconsistencias que encuentre el programa computacional.

2.8.5 Sistema de Control de la Fuerza de Agarre de la Mano de un Robot

En la figura 2.12 se puede ver el diagrama esquemático de un sistema de control de la fuerza de agarre, que utiliza un dispositivo sensor de fuerza y otro de deslizamiento. Si la fuerza de aprehensión es demasiado pequeña, la mano del robot dejará caer el objeto mecánico y si es demasiado grande, lo puede dañar o aplastar. La mano recoge y levanta el objeto con la fuerza de agarre preajustada. Si hay algún deslizamiento durante el ascenso, será detectado por el dispositivo censor de deslizamiento, el que enviará una señal de retorno al controlador, el cual a su vez aumentará la fuerza de agarre. De este modo, se puede lograr una fuerza razonable que evite el deslizamiento, pero que no produzca ningún daño al objeto.

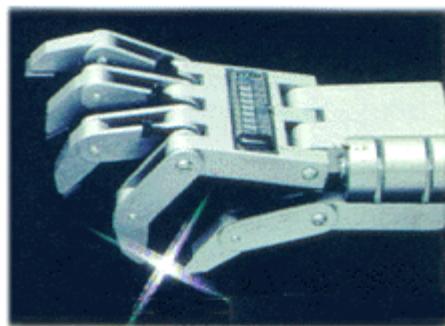


Figura 2.12: Fuerza de Agarre de la Mano de un Robot

Fuente: www.aurova.ua.es

Componentes:

- **Objeto:** Brazo robot.
- **Subsistema Controlador:** Sensor electromecánico de agarre.
- **Variable por controlar:** Presión sobre el objeto.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Sensor de medición de presión sobre el objeto.
- **Comparador:** Dispositivo que compare el valor de la presión sobre el objeto y el valor máximo permitido.
- **Reacción:** Apretar o aflojar.

2.8.6 Sistema de Control Hidráulico

En un sistema hidráulico el medio de transmisión de potencia es un fluido bajo presión. El fluido es generalmente aceite debido a que también actúa como lubricante. Los sistemas neumáticos trabajan de la misma forma que los sistemas hidráulicos con la excepción de que en este caso el medio utilizado es aire, lo que es una ventaja cuando el riesgo de fuego o explosión debe ser minimizado.

Un ejemplo de sistema mecánico hidráulico es el mecanismo de *power steering* en un automóvil (véase la figura 2.13).

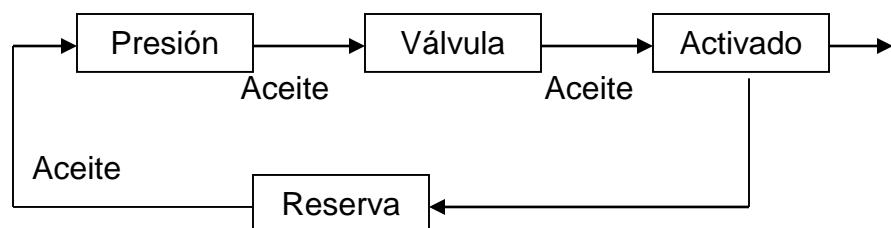


Figura 2.13: Sistema Hidráulico

Componentes:

- **Objeto:** Aceite.
- **Subsistema Controlador:** Válvula.
- **Variable por controlar:** Presión sobre el fluido.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Sensor de presión sobre el fluido.
- **Comparador:** Sistema que compara la presión y el valor máximo permitido.
- **Reacción:** Disminuir o aumentar la presión.

2.8.7 Sistema de Control de Calefacción

Los sistemas de control de calefacción son pocas veces como los hidráulicos. La mayoría de los sistemas de calefacción no tienen capacidad de enfriar. Si la temperatura excede el valor seleccionado, los sistemas de calefacción meramente se apagan y esperan a que el área se enfrié. Una recepción a esto es el sistema de control de clima usado en algunos automóviles, los cuales es el ciclo de acondicionamiento de calor y aire.

A continuación, presentamos un típico sistema de calefacción casero.

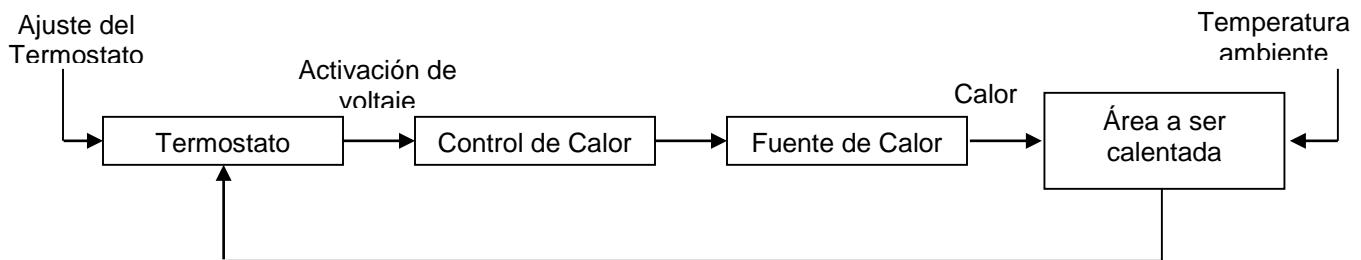


Figura 2.14: Control de Calefacción

Componentes:

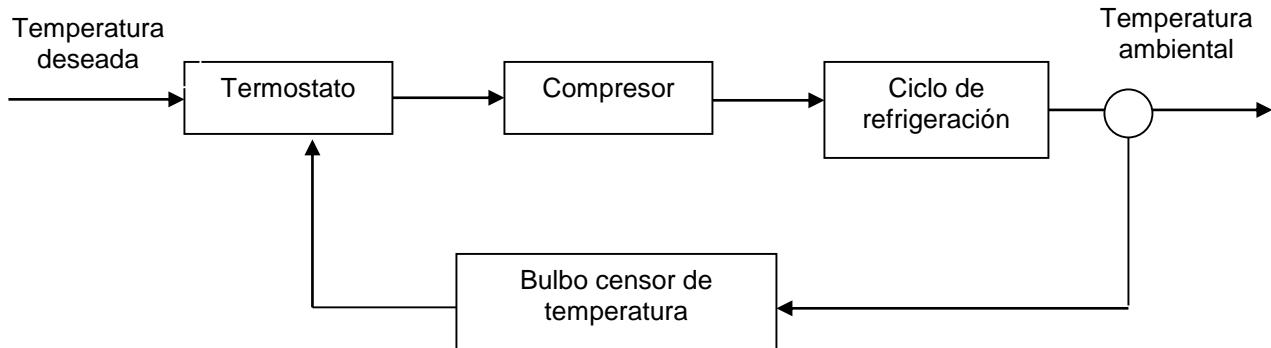
- **Objeto:** Sistema de calefacción.
- **Subsistema Controlador:** Control del calor.
- **Variable por controlar:** Temperatura del aire.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** El termostato que se encarga de sensar la temperatura del aire.
- **Comparador:** Dispositivo que compara la temperatura con el valor deseado.
- **Reacción:** Apagar o encender el sistema de calefacción.

2.8.8 Control de Temperatura de la Cabina de Pasajeros de un Coche

Imaginemos la dinámica de sistemas para este caso. La entrada al controlador es la temperatura deseada, convertida a un voltaje. La temperatura efectiva de la cabina se convierte a un voltaje por medio de un sensor, y se le retroalimenta al controlador para su comparación con la entrada. La temperatura ambiente y el calor del sol, transferido por radiación, actúan como perturbaciones debido a que no son constantes durante la marcha del vehículo. Este sistema utiliza tanto control retroalimentado como control de realimentado (el control de realimentación brinda acción correctiva antes que las perturbaciones afecten la salida).

La temperatura en la cabina del vehículo varía considerablemente, según el lugar en que se mida. En vez de instalar múltiples sensores para medir la temperatura, y obtener un promedio de los mismos, es más económico colocar un ventilador de succión en el sitio donde normalmente los pasajeros sienten la temperatura. Entonces la temperatura del aire del extractor brinda una indicación de la temperatura de la cabina y se le considera como salida del sistema.

El controlador recibe la señal de entrada, la señal de salida y las señales de sensores desde las fuentes de perturbación. A su vez, el controlador envía una señal óptima de control al equipo acondicionador de aire para controlar la cantidad de aire refrigerado, de modo que la temperatura sea igual a la temperatura deseada.



Componentes:

- **Objeto:** Cabina de pasajeros de un coche.
- **Subsistema Controlador:** El control retroalimentado
- **Variable por controlar:** Temperatura de la cabina de pasajeros de un coche.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Sensores que midan el valor de la temperatura de la cabina.
- **Comparador:** Dispositivo que compare la temperatura efectiva de la cabina de pasajeros con la temperatura deseada.
- **Reacción:** Aumentar o disminuir la cantidad de masa de refrigerante en el evaporador a fin de manejar la carga térmica y alcanzar el valor de la

temperatura deseada.

2.8.9 Sistema de Control de Tráfico

Como se ha indicado anteriormente, el control de tráfico por medio de señales activadas sobre una base de tiempo, constituye un sistema de control de ciclo abierto. Sin embargo, si la cantidad de automotores esperando en cada señal de tráfico en un área congestionada se mide continuamente, y esa información se lleva a una computadora central que controla tales señales, el sistema se convierte en sistema de ciclo cerrado.

El movimiento de tráfico en redes es muy complejo porque la variación en el volumen de tráfico depende mucho de la hora y el día de la semana, así como de muchos otros factores. En algunos casos se puede suponer una distribución de Poisson de llegadas a las intersecciones, pero esto no es necesariamente válido para todos los problemas de tráfico. De hecho, minimizar el tiempo de espera es un problema de control muy complejo.

Componentes:

- **Objeto:** Tráfico.
- **Subsistema Controlador:** Sistema de señales de tráfico.
- **Variable por controlar:** Flujo del tráfico.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Sensores que midan el movimiento del tráfico.
- **Comparador:** Dispositivo que mida la variación en el volumen de tráfico.
- **Reacción:** Transmitir las señales de tráfico necesarias para descongestionar el flujo automovilístico.

2.8.10 Sistema de Control de Inventario

Otro ejemplo de sistema de control de ciclo cerrado lo constituye la programación industrial del ritmo de producción y nivel de inventario. El nivel de inventario real, que es la salida del sistema, se compara con el nivel de inventario deseado que puede variar ocasionalmente según el mercado. Si aparece una diferencia entre el nivel de inventario real y el deseado, el ritmo de producción se ajusta de manera que la salida siempre iguale o esté cercana al valor deseado, que se elige para maximizar las utilidades.

Componentes:

- **Objeto:** Procesadores
- **Subsistema Controlador:** Sistema de Inventario
- **Variable por controlar:** Cantidad de procesadores en existencia
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Programa computacional que lee la cantidad existente cada vez que un vendedor vende y factura procesadores
- **Comparador:** Programa computacional que compara la cantidad en existencia

- con el punto de reorden de este artículo.
- **Reacción:** No hacer nada o solicitar que se compren más; ya que la existencia se encuentra por debajo del punto de reorden.

2.8.11 Sistemas Empresariales

Un sistema empresarial puede consistir de varios grupos. Cada tarea asignada a un grupo, representa un elemento dinámico del sistema. Para el funcionamiento correcto de tales sistemas, hay que establecer métodos retroalimentados para el control de los logros de cada grupo. El acoplamiento mutuo entre grupos funcionales debe ser mínimo para reducir atrasos inútiles en el sistema. Cuanto menor sea el acoplamiento más suave será el flujo de señales y de materiales.

Componentes:

- **Objeto:** Empresa
- **Subsistema Controlador:** Gerencia General y otras gerencias.
- **Variable por controlar:** Grupos de Trabajo.
- **Un observador o detector, que se encargue de revisar las señales que se reciban de la variable:** Asistente Ejecutivo de la Gerencia General.
- **Comparador:** Resultados de los grupos versus plan estratégico de la empresa.
- **Reacción:** Decisiones que toma el Gerente General para mejorar el acoplamiento de los grupos.

El capítulo 10 de este libro trata en mayor detalle la temática de la dinámica de una empresa. Este tema se ha convertido en algo muy crucial para la gerencia de una empresa en este nuevo siglo y se ha descubierto que la dinámica de sistemas puede aportar considerablemente al mejoramiento de la eficiencia.

2.9 Resumen

Toda actividad que implique el cumplimiento de una función para lograr un propósito, depende de un sistema de control. Los modelos matemáticos son el arma fundamental en el diseño de los sistemas de control, porque sirven para la elaboración de una teoría unificada de control.

La precisión dentro de un sistema de ciclo abierto depende en gran medida del criterio y estimación de su operador. Un sistema de ciclo cerrado puede diseñarse para tratar de obtener error cero (o casi cero).

En un sistema retroalimentado, la variable es controlada por la salida del sistema, la cual es constantemente revisada y corregida para llevarla al nivel deseado lo que puede producir variaciones en la salida del sistema manteniéndola en constante cambio.

El controlador, dentro de un sistema de ciclo cerrado, compara una salida deseada de la planta con la salida real y actúa en consecuencia para reducir las dos.

2.10 Revisión Histórica

El primer trabajo significativo en control automático fue el **regulador centrífugo** de James Watt para el control de la velocidad de una máquina de vapor en el siglo XVIII. En 1922 Minorsky trabajó en controles automáticos de dirección en barcos y mostró como se podría determinar la estabilidad a partir de las ecuaciones diferenciales que describen el sistema. En 1934, Hazen, que introdujo el término “**servomecanismo**” para los sistemas de control de posición, estudió el diseño de servomecanismos repetidores capaces de seguir estrechamente una entrada cambiante.



Figura 2.15: Regulador Centrífugo de James Watt

Al inicio de los años 40, los **métodos de respuesta de frecuencia** posibilitaron a los ingenieros, el diseño de sistemas de control retroalimentado lineal que satisfacían las necesidades de comportamiento. A finales de esta década e inicios de los años 50, se desarrolló completamente el **método del lugar de las raíces** en el diseño de sistemas de control.

Los métodos de respuesta de frecuencia y del lugar de las raíces, que son el corazón de la **Teoría de Control Clásica**, llevan a sistemas que son estables y que satisfacen un conjunto de requerimientos de funcionamiento. La teoría de control Clásica trata problemas de entrada y salida única, por tanto, se vuelve impotente ante sistemas de múltiples entradas y salidas. Desde 1960, se ha desarrollado la **Teoría de Control Moderna** para hacerle frente a las plantas modernas y las necesidades rigurosas en exactitud, peso y costo en aplicaciones militares, espaciales e industriales. Gracias a la disponibilidad de computadoras, se ha desarrollado en gran medida esta metodología y actualmente se están desarrollando aplicaciones de la Teoría de Control Moderna a campos no ingenieriles como la Biología, Economía, Medicina y Sociología, esperando resultados óptimos en un futuro.



Control PID, conceptos básicos. Nota técnica 10, rev b

Enero 2010

Este documento posee derechos de autor reservados, © ARIAN S.A. 2010
Las marcas comerciales referidas son propiedad de sus respectivos dueños.
Su contenido puede ser reproducido para fines educativos mientras no sea
alterado.

Presentación y objetivos.....	2
Variable de proceso, Set Point y Error.....	3
Control On/Off.....	4
PWM y tiempo de ciclo.....	6
Control Proporcional.....	8
Error estacionario	11
Control Proporcional Derivativo PD.....	14
Control Proporcional Integral PI.....	18
Control Proporcional Integral Derivativo PID.....	23
Mandos de enfriamiento en los PID.....	26

Presentación y objetivos.

El presente documento tiene por objetivos :

- 1- Introducir algunos elementos de teoría de control en forma gradual y lo más natural e intuitiva posible.
- 2- Familiarizar al lector con los términos usados en la programación de controladores PID modernos (banda proporcional, "reset automático",etc) de modo que pueda programar su instrumento entendiendo cada parámetro que introduce.
- 3- Establecer la relación entre los términos usados en terreno por los instrumentistas y su equivalencia en la teoría formal de control moderno.

Los objetivos 1 y 2 son los principales, a ellos se orienta la mayor parte del documento y puede lograrlos cualquier lector con conocimientos básicos de matemáticas, familiarizado un mínimo con instrumentos y máquinas industriales.

El mismo no necesita leer los recuadros "NIVEL 2", estos son apropiados al lector con conocimientos de cálculo diferencial (integral y derivada) y en algún momento estarán destinados a lectores con conocimientos de teoría de control y transformada de Laplace.

En todo caso estos son una parte mínima del texto y se incluyen solo en atención al objetivo 3.

ARIAN S.A.
El Comendador 2340, Providencia
Santiago, Chile
Fono: 56-2-4218333
web site: <http://www.arian.cl/>

Variable de proceso, Set Point y Error.

La variable medida que se desea estabilizar (controlar) recibe el nombre de variable de proceso ("Process Value") y se abrevia Pv.

Un buen ejemplo de variable de proceso es la temperatura, la cual mide el instrumento controlador mediante una termocupla o una Pt100.

Otro ejemplo de una Pv puede ser un caudal (litros/minuto), este se mide mediante algún sensor y su transductor adecuado que entreguen una señal de 4 a 20mA proporcional al caudal. La corriente 4 a 20mA entra al controlador en donde se reescal a las unidades originales de la Pv.

Ejemplos adicionales de Pv pueden ser velocidad, presión, humedad, etc cada una de las cuales se mide con el sensor apropiado y se convierten comúnmente a 4...20mA para ingresar al control.

En adelante durante todo este tutorial se usará la temperatura como ejemplo de Pv por ser lo más intuitivo.

Esto no hace perder generalidad, pues todo lo que se hable sobre temperatura y hornos es perfectamente análogo a sistemas de control de velocidad, presión, etc.

El valor prefijado Sp ("Set Point"), es el valor deseado de la variable de proceso, es el valor al cual el control se debe encargar de mantener la Pv.

Por ejemplo en un horno la temperatura actual es 155 °C y el controlador esta programado de modo de llevar la temperatura a 200°C.

Luego

$$\begin{aligned}Pv &= 155 \\Sp &= 200\end{aligned}$$

Se define error E como la diferencia entre la variable de proceso Pv y el set point Sp,

$$E = (Sp - Pv)$$

En el ejemplo anterior

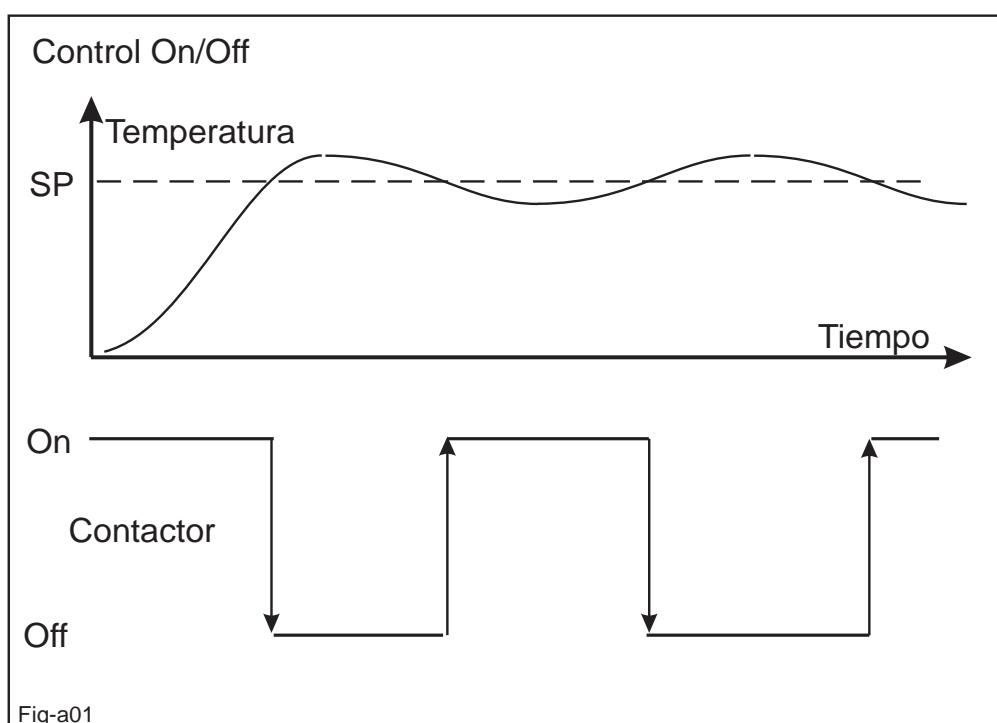
$$E = (Sp - Pv) = (200 - 155) = 45$$

Recuerde que el error será positivo cuando la temperatura sea menor que el set point, Pv < Sp .

Control On/Off.

Tomemos por ejemplo, el caso de un horno eléctrico. La temperatura aumenta al activar las resistencias calentadoras mediante un contactor, gobernado a su vez por un relé dentro del controlador.

El modo de control ON/OFF es lo más elemental y consiste en activar el mando de calentamiento cuando la temperatura está por debajo de la temperatura deseada Sp y luego desactivarlo cuando la temperatura esté por arriba.



Inevitablemente debido a la inercia térmica del horno y los tiempos de propagación internos, la temperatura estará continuamente fluctuando alrededor del Sp.

La inercia térmica es consecuencia de la masa caliente dentro del horno la que acumula calor y retarda su calentamiento y enfriamiento.

También colabora a la amplitud de las oscilaciones el retardo en la propagación del calor en el interior del horno desde la resistencia calentadora hasta el sensor de temperatura.

No es difícil imaginar que las fluctuaciones aumentarán cuanto mayor sea la inercia térmica del horno (retardo).

Evidentemente este algoritmo de control no es el más adecuado cuando se desea una temperatura constante y uniforme, pero tiene la ventaja de provocar poco desgaste en los contactores electromecánicos, pues estos se activan y desactivan lo mínimo necesario.

Incluso para reducir aún más el desgaste a veces se efectúa un control ON/OFF con histéresis, es decir que el mando de calentamiento se activa unos pocos grados por arriba del Sp y se desactiva unos grados por abajo del Sp, provocando por supuesto una mayor fluctuación de la temperatura pero menor funcionamiento de los contactores.

El control ON/OFF con histéresis se usa generalmente en cámaras frigoríficas donde la temperatura debe estar dentro de un rango y se desea activar y desactivar lo menos posible los motores del sistema de refrigeración.

Ejemplo en el software VirtualPID:

Ejecute VirtualPID.exe y realice los siguientes comandos:

- File/Open/ 1_OnOff.sim
- Start

PWM y tiempo de ciclo.

Para poder controlar la temperatura con menos fluctuaciones, se debe poder entregar al horno una potencia gradual, en la medida requerida para mantenerlo a la temperatura deseada .

En el ejemplo anterior del control On/Off, el relé del mando de calentamiento estará activado 100%, entregando el máximo de potencia al horno o bien desactivado sin entregar potencia 0%.

El controlador proporcional entrega una potencia que varía en forma gradual entre 0 y 100% según se requiera y en forma proporcional al error

$$E = (Sp - Pv)$$

Lo más natural para entregar una potencia que varia de 0% a 100% en forma continua seria usar un horno a petróleo o gas y que el control module la potencia mediante la llave de paso del combustible.

La llave cerrada seria 0% de potencia y la llave totalmente abierta seria 100%.

El problema es que una válvula motorizada (gobernada mediante una señal 4..20mA) es costosa y solo se amerita en aplicaciones que así lo requieran.

Afortunadamente es posible modular de 0% a 100% la potencia que recibe un horno eléctrico mediante el mismo contactor que se usaría para un control on/off.

La idea es modular el tiempo de activación del contactor durante un lapso de tiempo fijo t_c , llamado “tiempo de ciclo”, menor al tiempo característico de respuesta del horno de modo que el horno reciba finalmente un promedio de la potencia.

Para ilustrar esta técnica conocida como modulación por ancho de pulso PWM (Pulse Width Modulation) recurrimos al siguiente ejemplo:

Suponiendo que un horno tiene una gran inercia térmica tal que tarda varios segundos en reaccionar subiendo la temperatura.

Al aplicarle pulsos de calentamiento periódicamente cada 4 segundos, modulados en duración, el horno experimentará un calentamiento proporcional al promedio de tiempo que ha estado activado, sin que su temperatura siga las fluctuaciones de 4 segundos con que se aplica la potencia.

Supongamos que nuestro horno funciona con un calefactor de 1000W, si se requiere una potencia de 500W, equivalente a 50% del total, entonces el relé se activa 2 segundos y desactiva otros 2, para luego empezar otro ciclo.

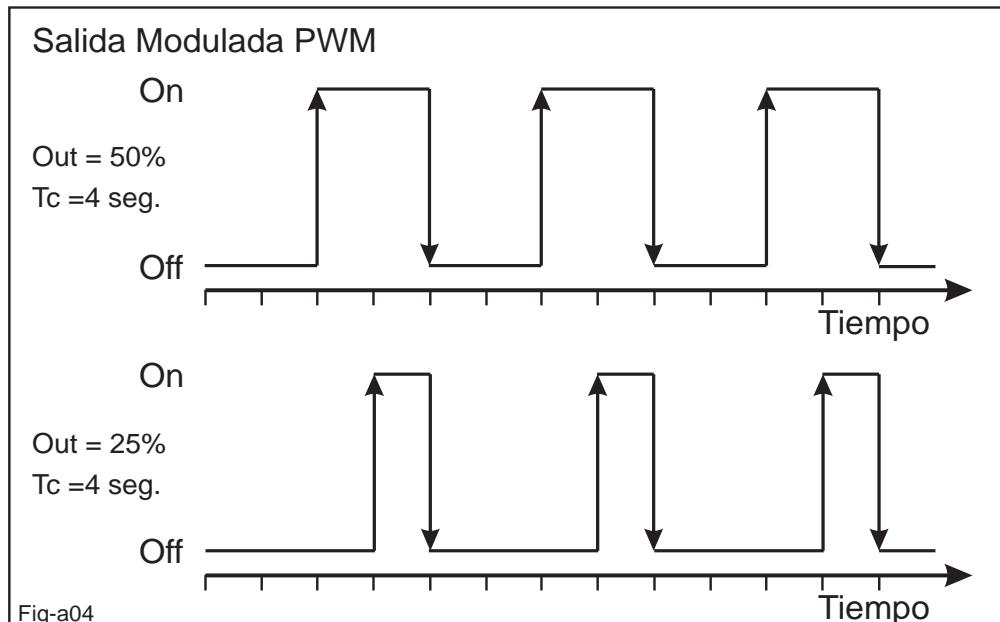
El efecto neto será que el horno recibe 50% de la potencia pero la temperatura no fluctúa al ritmo del “tiempo de ciclo” pues el horno no alcanza a responder a ciclos tan rápidos.

Siguiendo con el ejemplo, si hace falta 250W, es decir 25% de la potencia basta con tener el relé 1 segundo activado y 3 segundos desactivado.

Para sistemas típicos el “tiempo de ciclo” se ajusta entre 1 y 200 segundos según sea el caso.

A mayor “tiempo de ciclo”, menos desgaste de los contactores, pero siempre tiene que ser inferior al tiempo característico del sistema.

La práctica recomendada es usar un tiempo de ciclo inferior a la mitad del tiempo característico del sistema, el cual mas adelante se vera como determinar.



Ejemplo en el software VirtualPID:

Ejecute VirtualPID.exe y realice los siguientes comandos:

- File/Open/ 2_Pwm.sim
- Start
- En el formulario “Control” introduzca Output% = 50 y observe el comportamiento.

Control Proporcional.

Ya se menciono que el controlador proporcional entrega una potencia que varía en forma proporcional al error

$$E = (Sp - Pv)$$

Para poner en marcha un controlador proporcional se deben fijar los siguientes parámetros:

- la temperatura deseada Sp , por ej. Sp = 200 °C
- la banda proporcional Pb, por ej. Pb = 10 %.
- el tiempo de ciclo tc, por ej. tc = 4 seg.

Por supuesto no hace falta definir el tiempo de ciclo si se va a usar una válvula motorizada como mecanismo de control.

La banda proporcional Pb se programa en el controlador como un porcentaje del Sp. Corresponde a una banda de temperatura situada por debajo del Sp a lo largo de la cual, la potencia de salida variará proporcionalmente al error, disminuyendo cuanto más cercana sea la temperatura al Sp.

Internamente el controlador realizará el cálculo del porcentaje de salida "Out" mediante la siguiente fórmula:

$$Out = [100\% * \frac{E}{banda}]$$

con

$$E = (Sp - Pv)$$

$$banda = \frac{Pb}{100\%} * Sp$$

El paréntesis cuadrado [] se usa para indicar saturación, es decir que si al evaluar el interior, resulta mayor de 100%, se deja en 100% y si resulta 0 o negativo, se deja en 0%.

Para los valores del ejemplo Sp=200°C y Pb=10%, la potencia determinada por el control variará a lo largo 20°C abajo del Sp.

$$banda = (Pb * Sp / 100\%) = 10\% * 200C / 100\% = 20C$$

Es decir que la banda a lo largo de la cual variará gradualmente la potencia será: 180°C...200°C.

Por ejemplo si la temperatura del horno es igual o menor de 180°C, la salida de control (potencia) será 100%.

Cuando la temperatura esté en la mitad de la banda, es decir en 190°C la salida será 50% :

$$Out = [100\% * \frac{(200 - 190)}{(10 * 20/100)}] = 50\%$$

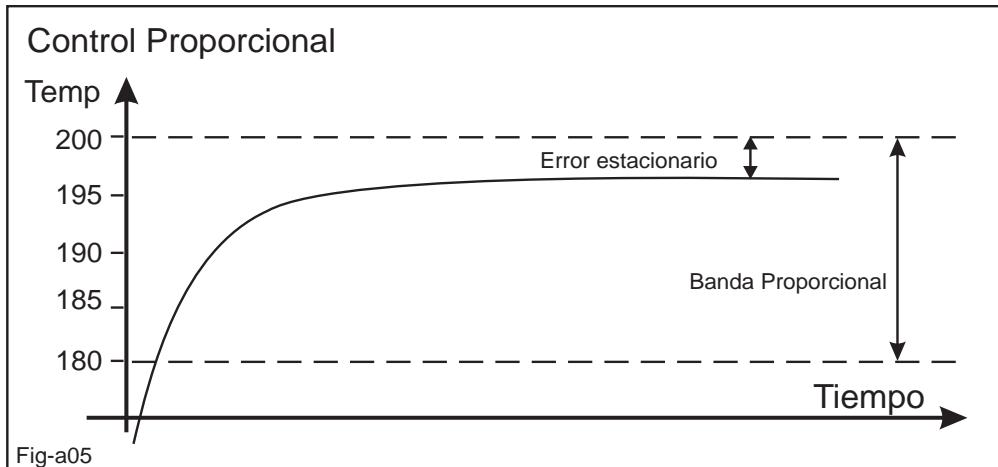
Al llegar la temperatura a 200 °C la salida será 0% :

$$Out = [100\% * \frac{(200 - 200)}{(10 * 20/100)}] = 0\%$$

Tabla 1		
Pv (Temperatura)	Error = Sp-Pv	Out=100%*E/20
C	C	%
210	-10	0
202	-2	0
200	0	0
199	1	5
198	2	10
195	5	25
190	10	50
185	15	75
182	18	90
181	19	95
180	20	100
170	30	100
110	90	100

En la Tabla 1 se observa la potencia de salida (Out%) suministrada por el controlador a distintas temperaturas para este ejemplo.

En el siguiente gráfico de temperatura vs. tiempo se observa el comportamiento típico de un control proporcional.



Es posible imaginar un control On/Off como uno proporcional con $P_b = 0\%$, pues cuando la temperatura esta arriba del Sp ($E < 0$) la salida es 0% y cuando la temperatura esta abajo del Sp ($E > 0$) es 100%.

Por lo tanto es importante tener presente que mientras menor sea la banda proporcional, el control proporcional se comportará más parecido al On/Off, es decir tenderá a presentar oscilaciones alrededor del Sp .

Ejemplo en el software VirtualPID:

Ejecute VirtualPID.exe y realice los siguientes comandos:

- File/Open/ 3_P.sim
- Start
- En el formulario “Sistema” presione el botón “Reset” y observe el comportamiento.

Error estacionario

El control proporcional presenta el problema que la temperatura jamás se estabilizará justo en el valor del Sp.

En la práctica se estaciona siempre en un punto dentro de la banda proporcional, produciendo así el "error estacionario".

La razón es fácil de entender recurriendo a nuestro ejemplo. Supongamos que la temperatura se estacionará en 200 °C en forma estable y permanente, entonces la salida sería 0%.

Pero siempre es necesario suministrarle al horno algo de potencia, por lo menos cómo para compensar las pérdidas de calor al medio ambiente o calor contenido en el material que sale del horno.

Es evidentemente imposible que el horno se mantenga a 200°C con los calefactores permanentemente apagados.

Luego la temperatura tiene que descender un poco, lo suficiente para que la potencia de suministrada por los calefactores sea igual a las pérdidas de energía del horno (energía que sale del horno).

Supongamos que para nuestro horno las pérdidas son tales que hace falta 25% de la potencia de los calefactores para mantener la temperatura alrededor del Sp.

Luego la temperatura se estacionará establemente a 195°C, pues para esa temperatura la salida es 25% (ver Tabla 1).

El valor del error en este punto, 5°C, se le llama error estacionario.

El error estacionario se puede reducir disminuyendo la banda proporcional. En nuestro ejemplo de horno usando ahora Pb = 5%, la temperatura para tener Out = 25% es 197.5°C.

$$Out = [100\% * \frac{E}{banda}] = [100\% * \frac{(Sp - Pv)}{(Pb * Sp / 100\%)}] =$$

$$Out = [100\% * \frac{(200 - 197.5)}{(5 * 200 / 100)}] = 25\%$$

Pero reducir mucho la banda proporcional volverá oscilatorio nuestro sistema (más parecido a un On/Off), luego existe un límite inferior para seguir reduciendo la banda proporcional y siempre habrá algo de error estacionario.

En particular en los hornos que poseen mucha inercia térmica (mucho tiempo de retardo) se pueden presentar oscilaciones de la temperatura que solamente se podrán eliminar aumentando la banda proporcional y con ella aumentará el error estacionario.

Otro problema generado al aumentar la banda proporcional para eliminar las oscilaciones, es que el control pierde efectividad para responder rápidamente a perturbaciones externas (variaciones de la carga del horno, apertura de una puerta, etc)

Para mejorar la respuesta a transientes del control se utiliza un control proporcional derivativo.

NIVEL 2

Internamente el controlador realiza el cálculo del porcentaje de salida "Out" mediante la siguiente fórmula

$$Out = [100\% * \frac{E}{banda}]$$

$$E = (Sp - Pv)$$

$$banda = \frac{Pb}{100\%} * Sp$$

El paréntesis [] se usa para indicar saturación, es decir que si al evaluar el interior, resulta mayor de 100%, se deja en 100% y si resulta negativo, se deja en 0%.

La misma fórmula se puede reescribir en términos de la ganancia proporcional como:

$$Out = [100\% * Kp * E]$$

$$Kp = \frac{1}{banda} = \frac{100\%}{Pb} * \frac{1}{Sp}$$

Donde la constante Kp se conoce como ganancia proporcional del control y es inversamente proporcional a Pb .

En el control On/Off, se podría decir que Kp sería muy grande, tendiendo a infinito.

Ejercicios:

- 1) Un control proporcional tiene programado:

$$Sp = 500$$

$$Pb = 5\%.$$

Si $Pv = 490$ cual será el valor de la salida de control.?

Respuesta: Out = 40%.

- 2) Un control proporcional de temperatura tiene programado:
 $Sp = 300^{\circ}\text{C}$
 $Pb = 20\%$.
A que temperatura la salida de control será 50%.?

Respuesta: $Pv = 270^{\circ}\text{C}$.

- 3) Un horno con control proporcional de temperatura tiene programado:
 $Sp = 800^{\circ}\text{C}$
 $Pb = 10\%$.
El sistema está estable con una temperatura pareja de 760°C .
- a) Cual es el error estacionario.?
b) Al colocar $Pb = 5\%$, cual será la temperatura aproximada a la que estabilizará el horno. ?

Respuesta: a) 40°C , b) 780°C

Ejemplo en el software VirtualPID:

- File/Open/ 3_P.sim
- Start
- En el formulario “Sistema” presione el botón “Reset” y observe que después de un tiempo el error estacionario se estabiliza en 13.3°C .
- En el formulario “Control” aumente Pb a 20% y observe como cambia el error estacionario.

Control Proporcional Derivativo PD.

Un control PD es uno proporcional al que se le agrega la capacidad de considerar también la velocidad de la temperatura en el tiempo.

De esta forma se puede "adelantar" la acción de control del mando de salida para obtener así una temperatura más estable.

Si la temperatura esta por debajo del Sp, pero subiendo muy rápidamente y se va a pasar de largo el Sp, entonces el control se adelanta y disminuye la potencia de los calefactores.

Al revés si la temperatura es mayor que el Sp, la salida debería ser 0% pero si el control estima que la temperatura cae muy rápido y se va pasar para abajo del Sp, entonces coloca algo de potencia para ir frenando el descenso brusco.

La acción derivativa es llamada a veces "rate action" por algunos fabricantes de controles porque considera la "razón de cambio" de la temperatura.

Para entender a fondo el control PD usaremos el mismo ejemplo anterior del horno pero agregamos ahora un nuevo parámetro llamado constante derivativa D, medido en segundos.

Internamente el controlador realizará ahora el cálculo:

$$Out = [100\% * \frac{(E - D * Vel)}{banda}]$$

$$banda = (Pb / 100\%) * Sp$$

Donde "Vel" es la velocidad de la temperatura medida por el controlador, en °C/seg.

Para este ejemplo fijamos D = 5 seg. Igual que antes Sp=200 °C y Pb=10%. Supongamos que en un momento dado, la temperatura del horno es de 185°C y está subiendo a una velocidad Vel = 2 °C/Seg..

En un control proporcional la salida debería ser de 75%.

$$Out = [100\% * \frac{(E)}{banda}] = [100\% * \frac{15C}{20C}] = 75\%$$

Pero en este caso el control PD toma en cuenta la velocidad de ascenso de la temperatura y la multiplica por la constante derivativa D y obtiene :

$$Out = [100\% * \frac{(E - D * Vel)}{banda}] = [100\% * \frac{15C - 5seg * 2C / seg}{20C}] = 25\%$$

entonces a pesar que la temperatura actual es 185 °C, la salida es 25% en vez de 75%, al considerar la velocidad de ascenso de la temperatura.

La acción derivativa ocurre también fuera de la banda proporcional, en la Tabla 2 se puede ver que para esta misma velocidad de ascenso de 2 °C/seg, la salida deja de ser 100% a partir de 170°C.

$$Out = [100\% * \frac{(E - D * Vel)}{banda}] = [100\% * \frac{(200 - 170)C - 5seg * 2C/seg}{20C}]$$

$$Out = [100\% * \frac{20C}{20C}] = 100\%$$

De la misma forma si la temperatura está sobre 200 °C pero descendiendo rápidamente, (velocidad negativa) por ejemplo -1°C/seg, entonces el control activará antes y con mayor potencia la salida intentando que no baje de 200 °C.

En la tabla 2 se observan las potencias a distintas temperaturas para 3 casos: control proporcional, control PD con velocidad positiva y el mismo control PD con una velocidad negativa.

El control PD permite obtener un control mas estable, sin oscilaciones y sin necesidad de sacrificar la respuesta ante perturbaciones aumentando la banda proporcional.

TABLA 2 SP = 200°C , Pb = 10%			
Control	P	PD	PD
		D = 5seg	D = 5 seg
		Vel = 2 C/seg	Vel = -1C/seg
	$[100\% * \frac{E}{20}]$	$[100\% * \frac{E - 5 * 2 / seg}{20}]$	$[100\% * \frac{E + 5 * 1 / seg}{20}]$
Pv	Out	Out	Out
204	0%	0%	5%
201	0%	0%	20%
200	0%	0%	25%
199	5%	0%	30%
198	10%	0%	35%
195	25%	0%	50%
190	50%	0%	75%
185	75%	25%	100%
182	90%	40%	100%
181	95%	45%	100%
180	100%	50%	100%
171	100%	95%	100%
170	100%	100%	100%

NIVEL 2

Dado que el Sp es constante, la velocidad del error es:

$$\frac{dE}{dt} = \frac{d(Sp - Pv)}{dt} = -\frac{dPv}{dt}$$

Luego la ecuación para el control PD se puede escribir:

$$Out = [100\% \cdot Kp \cdot (E + Kd \cdot \frac{dE}{dt})]$$

Donde $Kd = D$ es la ganancia derivativa y

$$Kp = \frac{100\%}{Pb} \frac{1}{Sp}$$

Se puede retirar los paréntesis [] y el 100%, sin olvidar que en la practica hay saturación de la salida.

$$Out = Kp \cdot (E + Kd \cdot \frac{dE}{dt})$$

Ejercicios:

- 1) Un control PD tiene programado:
 $Sp = 500$,
 $Pb = 5\%$
 $D = 10 \text{ seg.}$
Si Pv va en aumento de forma que incrementa 30 unidades cada minuto.
En el instante que Pv pase por 490, cual será el valor de la salida de control.?
Respuesta: $Out = 20\%$.

- 2) Un control PD tiene programado:
 $Sp=500$,
 $Pb=5\%$
 $D=10 \text{ seg.}$
Si Pv es superior a Sp y va en descenso 90 unidades cada minuto.
En el instante que Pv sea 505, cual será el valor de la salida de control.?
Respuesta: $Out=40\%$.

Ejemplo en el software VirtualPID:

- File/Open/ 4_PD.sim
- Start
- En el formulario “Sistema” presione el botón “Reset”
- En el formulario “Control” varíe el valor del parámetro D en el rango 0 a 10 y observe el comportamiento.

Control Proporcional Integral PI.

Retomaremos ahora el problema inconcluso del "error estacionario" tratado en la sección dedicada al control proporcional.

Para ello continuaremos utilizando el mismo ejemplo de nuestro horno a 200°C con Pb=10%, estacionado a 195°C para entregar 25% de potencia.

Una forma de eliminar el error estacionario podría ser aumentando, mediante algún ajuste manual, un 25% la salida del control de modo que se estacione en 200°C.

Algunos controles antiguos permitían este ajuste y era llamado "Manual reset" o "Ajuste manual", el problema con este ajuste es que será efectivo mientras el horno esté con las mismas características que cuando se realizó.

Por ejemplo, si por alguna razón las pérdidas del horno disminuyen a 20%, la temperatura subirá por arriba de 200°C creando un error por arriba del Sp.

La forma efectiva de solucionar el problema del error estacionario es agregando al control proporcional el término "Integral" llamado también a veces "automatic reset" o "reset action", nosotros lo llamaremos "acción integral".

El control será el mismo proporcional, pero a la salida se le suma la acción integral, que la corrige tomando en cuenta la magnitud del error y el tiempo que este ha permanecido.

Para ello se le programa al control una constante *I*, que es formalmente "la cantidad de veces que aumenta la acción proporcional por segundo".

La salida se corrige en una cantidad equivalente a la integral del error multiplicada por *I*.

Si parece complicado, es porque en realidad lo es un poco. Recurriendo a una analogía de un "saco" se podrá entender cómo opera la acción integral. La integral del error es simplemente la suma acumulada de los errores medidos por el control cada segundo.

Es como un saco al cual se le va metiendo o acumulando cada segundo una cantidad equivalente al error medido en ese segundo.

Si existe un error de 5°C, entonces el saco va creciendo, aumentando su tamaño cada segundo en una cantidad de 5.

Si el error es 0°C, entonces nada se le agrega al saco y este permanece en el mismo tamaño.

Pero si existe un error negativo, por ejemplo T=202 °C, entonces se le sacará 2 al saco cada segundo que permanezca este error y el tamaño del saco irá disminuyendo.

La idea es que la salida del control será la misma que en un control proporcional mas la magnitud del saco multiplicada por la constante **I**

Internamente el control PI realiza el cálculo con la siguiente fórmula:

$$Out = [100\% * \frac{E + I \cdot (EAcumulado)}{banda}]$$

$$E = (Sp - Pv)$$

$$banda = (Pb / 100\%) * Sp$$

Donde EAcumulado es el error acumulado o “integral del error” o simplemente nuestro saco.

Veamos en la Tabla 3 que ocurre al horno con una acción integral

$$I = 0.04/\text{seg.}$$

$$Sp=200^\circ\text{C},$$

$$Pb=10\%.$$

TABLA 3 SP = 200°C , Pb = 10%, I = 0.04seg				
Tiempo	Pv	E	EAcumulado	$[100\% * \frac{E + I \cdot EAcumulado}{20}]$
Seg	C	C	C	%
0	195	5	0	25%
1	195	5	5	26%
2	195	5	10	27%
3	195	5	15	28%
4	195	5	20	29%
5	196	5	25	30%
6	196	4	29	25.8%
7	196	4	33	26.6%
8	196	4	37	27.4%
9	196	4	41	28.2%
10	196	4	45	29%
11	197	3	48	24.6%
12	197	3	51	25.2%
13	197	3	54	25.8%
14	197	3	57	26.4%
15	198	2	59	21.8%
16	198	2	61	22.2%
17	198	2	63	22.6%
....	
Xx	200	0	125	25%

Con la temperatura estacionada en 195°C (por efecto de las pérdidas estimadas en 25%) a partir del tiempo 1, se activa la acción integral, desde ese momento cada segundo el saco aumenta en 5°C.

En ese primer segundo el acumulado es 0 y se le agrega 5C con lo que la potencia de salida aumentará 1%, de 25% a 26% .

$$Out = [100\% * \frac{E + I \cdot (EAcumulado)}{banda}] = [100\% * \frac{5 + 0.04 \cdot (5 + 0)}{20}] = 26\%$$

En el segundo 6 el aumento de potencia de salida habrá forzado el alza de la temperatura y así la reducción del error a 4°C.

El alza de la potencia de salida continua pero más lenta, ya que el Error es menor .

El proceso continua de la misma forma aumentando el tamaño del saco hasta que en algún momento el error sea definitivamente eliminado, pues solo en ese momento se detendrá el incremento de la potencia de salida.

Entonces será el saco (la integral) quien supla el "ajuste" de la potencia necesaria para mantener el horno a 200°C.

Por muy pequeño que sea el valor programado de **I**, siempre corregirá el error estacionario, pero tardará más tiempo en hacerlo.

Al revés si se programa un valor excesivo de **I**, entonces la acción integral tendrá mucha fuerza en la salida y el sistema alcanzará rápidamente el Sp, pero lo más probable es que siga de largo por efectos de la inercia térmica.

Entonces la acción integral (con error negativo) será en sentido contrario, el saco irá disminuyendo rápidamente con la magnitud del error.

Como consecuencia habrá una excesiva disminución de la potencia de salida y la temperatura probablemente baje del Sp, entrando así el sistema en un ciclo oscilatorio.

En la práctica normalmente **I** deberá ser grande solo en sistemas que reaccionan rápidamente, (por ejemplo controles de velocidad de motores) y pequeño para sistemas lentos con mucha inercia. (Por ejemplo hornos)

En general los valores de la constante **I** son relativamente pequeños, para la mayoría de los sistemas el valor adecuado de **I** varía entre 0 y 0.08 1/seg..

Por ese motivo en los controles Arian el parámetro **I** se programa multiplicado por 10000.

Es decir que para tener una acción integral de 0.01/seg, se programa

$$I = 0.01 * 10000 = 100.$$

Ejercicios:

- 1) Revisar los cálculos presentados en la tabla 3.

Ejemplo en el software VirtualPID:

-File/Open/ 5_PID.sim

-Start

-En el formulario “Sistema” presione el botón “Reset”

-En el formulario “Control” varíe el Sp y compruebe que siempre desaparece el error estacionario.

NIVEL 2

Internamente los controladores Arian realizan el cálculo con la siguiente fórmula:

$$Out = Kp \cdot (E + Ki \cdot \int Edt + Kd \cdot \frac{dE}{dt})$$

$$Kp = \frac{100\%}{Pb} \frac{1}{Sp}$$

$$Ki = \frac{I}{10000}$$

$$Kd = D$$

Donde $\int Edt$ es la integral del error en el tiempo (lo mismo que el saco de nuestra analogía) y Sp, Pb, D, I son los números introducidos por el teclado del instrumento.

La constante Ki es la ganancia integral .

El efecto del término integral es más beneficioso que un simple “reset automático”, pues el sistema de control se convierte en uno de segundo orden en vez de primer orden.

Esto implica una mejor respuesta a transientes y mejor rendimiento en general que un control proporcional con potencia de salida ajustada manualmente.

Internamente del control el término integral $\int Edt$ está también restringido al rango de valores (saturación de la integral)

$$-Pb*10000/I \dots +Pb*10000/I$$

de modo que su aporte en la salida esté limitado al que podría realizar el error de una banda proporcional.

El motivo principal es evitar el “over shot” en la puesta en marcha del sistema. Es decir previene un aumento excesivo de la parte integral cuando en la puesta en marcha el Error es grande durante un largo tiempo.

Control Proporcional Integral Derivativo PID.

A estas alturas el lector ya debe intuir que un control PID es un controlador proporcional con acción derivativa y acción integral superpuestas.

Así mismo el lector ya debe estarse preguntando cómo elegir los valores de los parámetros P_b , D , I , tc que debe introducir en su controlador PID.

Existe un conjunto de valores P_b , D , I que darán el rendimiento óptimo y encontrarlos requiere: conocimientos teóricos, habilidad obtenida mediante la experiencia y suerte.

Pero en la práctica cualquier conjunto de valores cercano al óptimo brindará un rendimiento aceptable y probablemente casi igual al óptimo.

Por este motivo los controles ARIAN admiten estos parámetros no en forma continua, si no como 16 valores discretos para cada uno.

Además afortunadamente existen un par de métodos experimentales para encontrar una aproximación de estos parámetros. (los que están descritos en el documentos)

Se dice que un sistema y su controlador tienen un comportamiento inestable cuando después de un tiempo razonable de funcionamiento y sin ocurrir perturbaciones externas, la temperatura permanece fluctuando en forma oscilatoria ya sea con un período regular o errático.

Este sería por ejemplo el caso de un control On/Off o uno proporcional con P_b muy chico.

Por otra parte un comportamiento estable es tal que la temperatura se mantiene en un valor constante mientras no ocurran perturbaciones externas.

Siempre se busca que el sistema de control sea estable, pero además, dentro de las condiciones de estabilidad existen 3 tipos de comportamiento bien definidos: control sub amortiguado, control con amortiguamiento crítico y control sobreamortiguado.

En las figura se muestran ejemplos de estos comportamientos. En cada uno de estos casos varía la velocidad de respuesta del sistema ante perturbaciones y a la vez la propensión a tener comportamiento inestable u oscilatorio.

El sistema sobreamortiguado tiene una velocidad de respuesta lenta, después que ocurra una perturbación, el sistema puede tardar en volver al Sp , pero la ventaja es que el sistema es muy estable y no adquiere comportamientos oscilatorios indeseables.

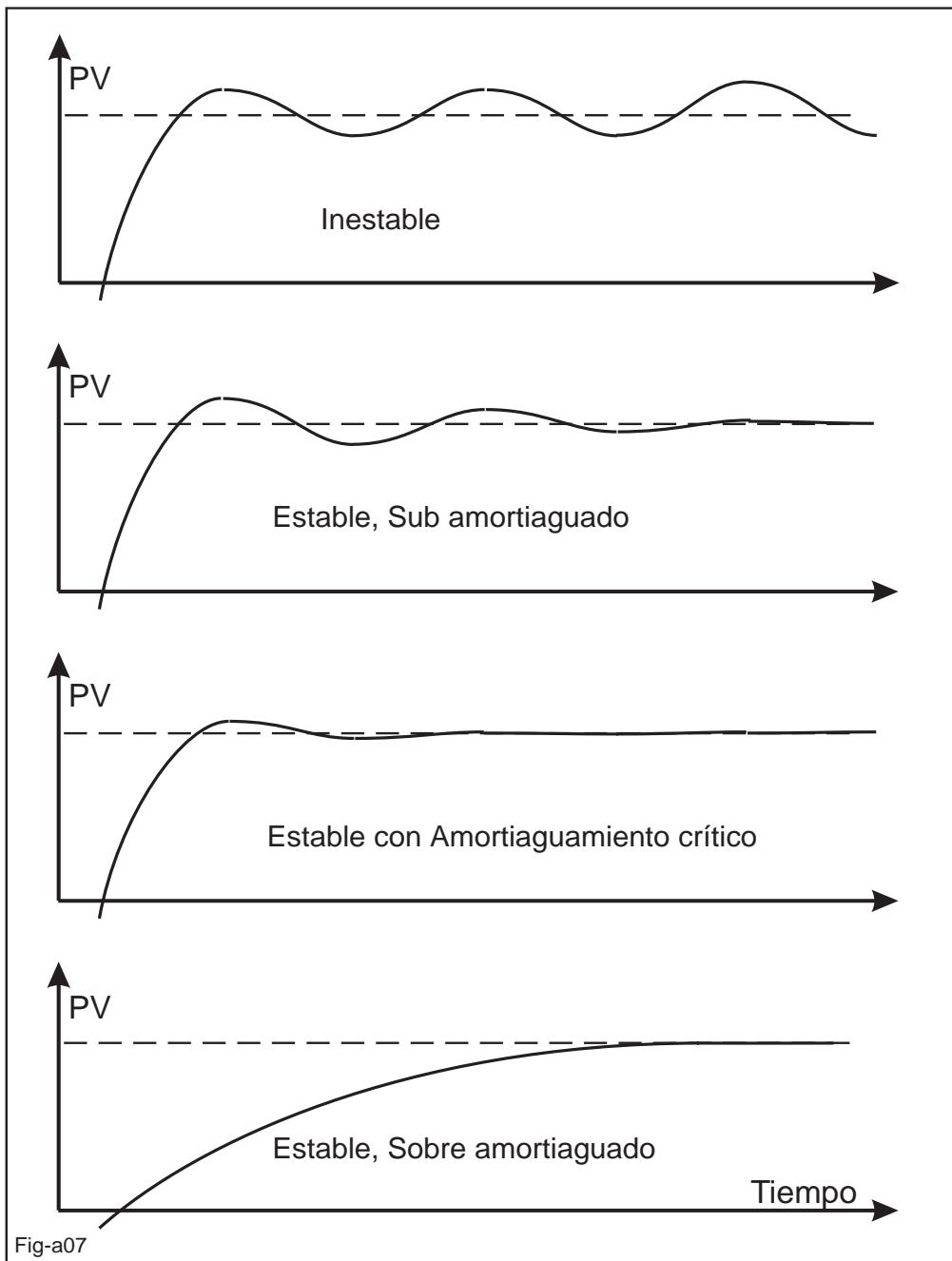
Esta condición tiende a ocurrir cuando la banda proporcional P_b es más grande de lo necesario. También puede deberse a una constante derivativa D muy

grandes, basta acordarse que la acción derivativa tiende a frenar la temperatura.

En el otro extremo, cuando un sistema se comporta de modo subamortiguado la velocidad de respuesta es muy buena pero pueden ocurrir varias oscilaciones de cierta amplitud antes que la temperatura llegue a un valor estable.

Suele ocurrir esta condición cuando la banda proporcional P_b es pequeña, la constante derivativa D chica y la constante de integración I grande.

El justo medio entre las condiciones anteriores es el amortiguamiento crítico. A esta condición corresponde los valores óptimos de los parámetros P_b , D , I . En este caso el sistema es bastante estable y la velocidad de respuesta es la mejor que se puede lograr.



Mandos de enfriamiento en los PID.

En ocasiones para poder controlar la temperatura de algunos sistemas se requiere enfriamiento forzado.

Un ejemplo común son las extrusoras de mangueras plásticas, en ellas es necesario calentar el material mediante calefactores eléctricos.

Pero al llegar a la temperatura de operación, en algunas zonas la temperatura sube por arriba S_p , efecto del roce mecánico ejercido por el tornillo que empuja el material.

Para bajar la temperatura se acostumbra en este caso a usar ventiladores eléctricos

Otro buen ejemplo podría ser una reacción química exotérmica entre reactivos que se están mezclando. Para iniciar la reacción hace falta subir la temperatura, pero una vez iniciada, la temperatura sube sola y en exceso. En este caso se acostumbra a usar circuitos de agua fría como medio de enfriamiento.

Es posible usar un control PID para ambos mandos (calentamiento y enfriamiento), esto se conoce en la industria como "dual PID" o "PID de rango expandido" o "PID + dPID" o algo así.

El control del mando de enfriamiento será PID también, usará los mismos valores de la integral y derivada del mando de calentamiento pero escalados al valor de una nueva "banda proporcional de enfriamiento" Pb_2 distinta a la banda proporcional del primer mando Pb .

La razón de esto es que las potencias o capacidades de calentamiento y enfriamiento son distintas y necesitan equilibrarse mediante estas bandas proporcionales.

Es natural que un sistema de agua, enfriá violentamente algo que costó mucho calentar mediante calefactores eléctricos.

Por ejemplo, un calefactor entrega al activarse 1000 Kw y el sistema ventilador puede retirar 2000Kw, si la banda proporcional Pb es 5 % entonces la banda proporcional de enfriamiento Pb_2 debería programarse en 10% para equilibrar las potencias.

$$1000\text{Kw}/5\% = 2000\text{Kw}/10\%$$

Una forma de estimar experimentalmente la relación entre las potencias de enfriamiento y calentamiento del sistema podría ser un experimento como el siguiente:

En un horno a 100°C se activa totalmente el mando de calentamiento, con lo cual la temperatura sube. Se observa que la temperatura llega a subir a una velocidad de 20°C / minuto.

Se desactiva el mando de calentamiento, se espera a que el sistema se estabilice un poco.

Se activa ahora el mando de enfriamiento y se observa que la temperatura desciende 30°C / minuto.

La idea es medir el cociente entre las velocidades calentamiento y enfriamiento pues este cociente es similar al de las potencias.

Del experimento se deduce que la banda proporcional Pb2 deberá ser 1.5 veces la banda proporcional Pb.

$$Pb2/Pb = 30^\circ\text{C}/20^\circ\text{C} = 1.5$$

En los controles PID + dPID se usa siempre una "**banda muerta**" o "**dead band**" o "separación del set point 2" dSP2, esto es una banda sobre el Sp en donde no se realiza acción de control de enfriamiento o calentamiento.

Es literalmente una banda muerta de unos pocos grados en donde la temperatura puede estar libremente sin que los mandos de control actúen.

Exactamente lo que no actúa es la acción proporcional, pero si lo hace las acciones derivativas e integral.

NIVEL 2

El cálculo interno de control para el PID - dPID será el siguiente:

$$Out = \left[\frac{100\%}{(Pb \cdot Sp)} \left(E + \frac{I}{10000} \int Edt + D \cdot \frac{dE}{dt} \right) \right]$$

$$Out2 = - \left[\frac{100\%}{(Pb_2 \cdot Sp_2)} \left(E_2 + \frac{I}{10000} \int Edt + D \cdot \frac{dE}{dt} \right) \right]$$

$$Sp_2 = Sp + dSp_2$$

$$E_2 = Sp_2 - Pv$$

Donde dSp_2 es la banda muerta y Pb_2 la banda proporcional del mando de enfriamiento.

Los valores de Out y Out2 varían en el rango 0...100% cada uno con saturaciones sobreentendidas en los cálculos.

Capítulo 3

Definición y Construcción de Modelos

Capítulo 3

Definición y Construcción de Modelos

Los objetivos de este capítulo son los siguientes:

- Presentar la definición de modelo.
- Presentar los diversos usos de los modelos.
- Presentar las diferentes clasificaciones de los modelos.
- Definir los pasos necesarios para la construcción de un modelo.
- Enumerar los diferentes elementos que conforman la estructura de un modelo.

3.1 Introducción

Uno de los objetivos de la simulación, como observamos en el capítulo anterior, es realizar experimentos sin tener la necesidad de alterar o perturbar el sistema real, ya sea por razones económicas o como protección al sistema. Este tipo de estudio pretende llegar a predecir cómo será el comportamiento futuro del sistema, en función de los cambios de ciertos parámetros. También es cierto que en la mayoría de los casos no se pueden cambiar deliberadamente los parámetros de un sistema real debido a que podrían causar errores graves de funcionamiento.

Al realizar estudios con un sistema económico, por ejemplo, no es factible cambiar arbitrariamente la oferta y la demanda del producto (tratándose del sistema real). Aunque, en ocasiones se da el caso, de que una empresa experimenta con diferentes estrategias de publicidad o mercadotecnia para luego comparar los resultados obtenidos. En el sistema real, se tiene muy poco o ningún control porque no es posible mantener constantes todas las variables. Si se lograra algún cambio, en los parámetros del sistema real, escaparía de nuestras manos la capacidad de predecir de manera razonable su futuro comportamiento.

Para que la simulación logre un comportamiento semejante al sistema real, sin los problemas descritos, se necesita construir un modelo. A continuación, se define el concepto de *modelo* para pasar inmediatamente a sus usos.

3.2 Definición de Modelo

Se pueden encontrar diferentes definiciones del concepto de *modelo*, a continuación, mencionaremos algunas de ellas:

- Representación física o abstracta de un sistema (grupo de objetos o ideas) que de alguna manera es diferente a su forma original.
- Conjunto de información recogida sobre un sistema con el propósito de estudiarla.

- Descripción detallada de un sistema.
- Abstracción de las propiedades dominantes de un sistema.
- Imagen o representación de los aspectos del sistema que se encuentran bajo estudio.

Analicemos estas definiciones. El término abstracción significa separación mental de todos los objetos o parte de las cualidades de un sistema, para su mejor estudio. Describir es representar gráficamente o por medio del lenguaje, las circunstancias que sirven para dar una idea clara de un sistema. Podemos condensar estas ideas y decir entonces que un modelo es:

Modelo es la descripción o representación de la abstracción de un sistema real, que puede ayudarnos a explicar, entender, predecir o mejorar el sistema bajo estudio.

3.3 Usos de los Modelos

En el capítulo anterior mencionamos que los estudios de simulación sirven como apoyo en el análisis y diseño de sistemas de cualquier tipo. Esto quiere decir, que los modelos, los cuales son herramientas de simulación, tendrán los mismos usos. Además, permiten diseñar procedimientos del sistema y posteriormente evaluarlos. Frecuentemente se reconocen cinco usos comunes y legítimos de los modelos:

3.3.1 Una Ayuda al Pensamiento

Esto significa que es un recurso explicativo que permite definir un problema o sistema, ya que muestra los pasos necesarios, y la secuencia de estos, para poder llegar a la solución. Por ejemplo, al construir una red PERT¹ de un determinado proyecto, se están indicando los pasos necesarios para su adecuada realización y terminación (ver Figura 3.1).

¹ Program Evaluation and Review Technique.

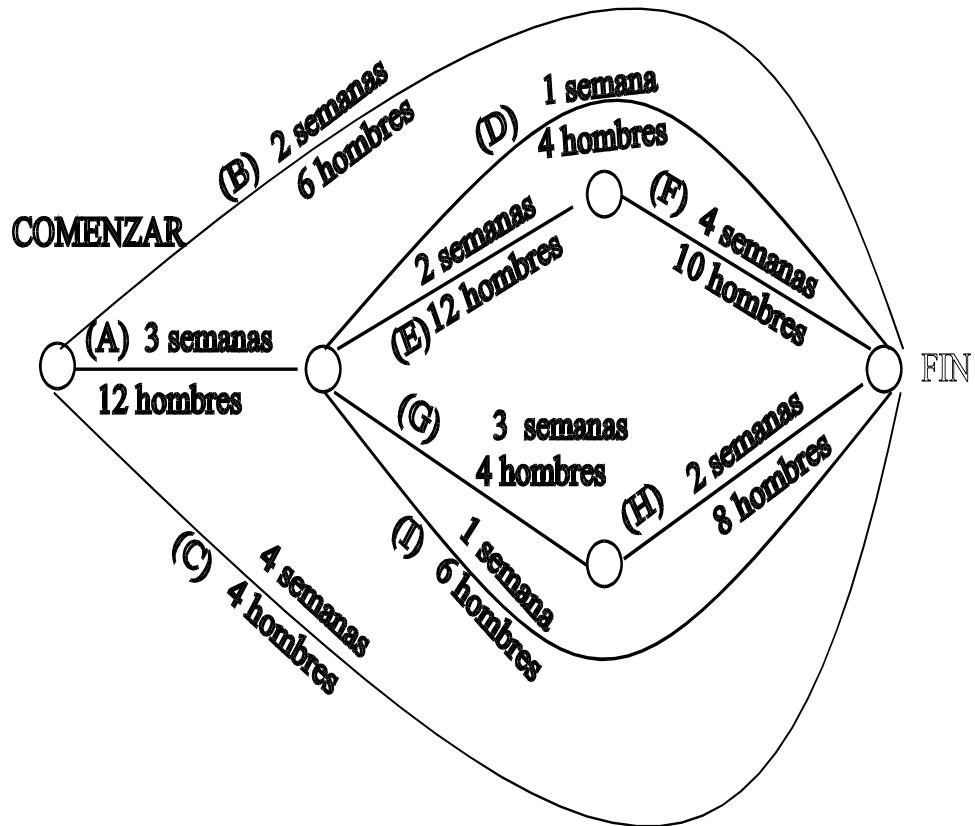


Figura 3.1: Ejemplo de un PERT

La red muestra las diferentes actividades y su secuencia, los tiempos de duración y los recursos que se requieren para cada actividad.

3.3.2 Una Ayuda a la Comunicación

Proporciona un modo de comunicación más eficiente y efectivo, permitiendo eliminar las ambigüedades que aparecen en un lenguaje verbal cuando se transmiten ideas o descripciones complejas. Hace más comprensible las relaciones causa y efecto (ver Figura 3.2).

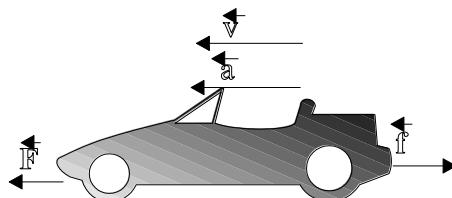


Figura 3.2: Ejemplo de un Modelo Visual

3.3.3 Una Ayuda al Entrenamiento e Instrucción

Los modelos son ideales para entrenar a las personas porque permiten:

- Practicar con el ambiente.
- Experimentar situaciones de crisis.
- Enfrentarse a eventualidades antes que estas ocurran.

Ejemplos:

- Vehículos espaciales para entrenar astronautas.
- Modelos para enseñar a conducir automóviles.
- Modelo de defensa aérea²: El propósito principal de éste modelo es el de proporcionar ayuda en la evaluación de ciertos parámetros de dichos sistemas defensivos en contra de ataques a poca altura por parte del avión enemigo. Mientras, el objetivo de la defensiva, entonces, es tratar de ocasionarle un alto costo a la ofensiva. El modelo presentado puede adaptarse con facilidad a situaciones más complejas y así poder observar su comportamiento.

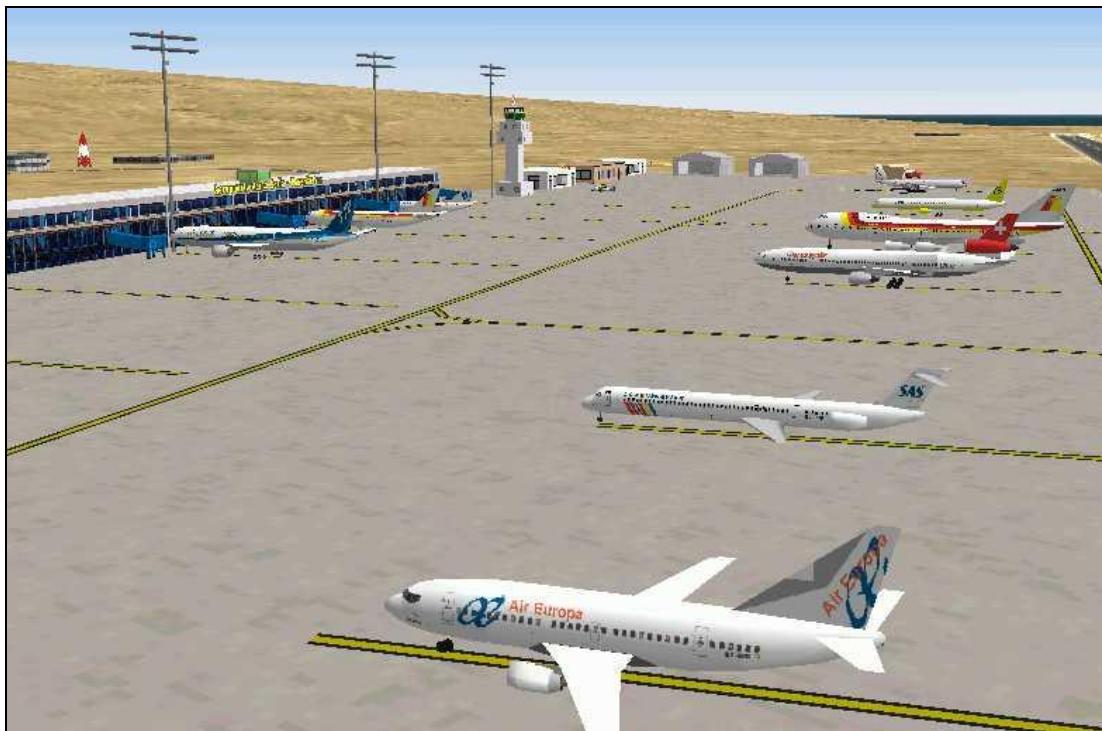


Figura 3.3: Ejemplo de Ayuda al Entrenamiento e Instrucción

² Robert Shannon, "Simulación de Sistemas", pp. 320-329.

3.3.4 Una Herramienta de Predicción y Ayuda al Desarrollo

El modelo ayuda a indicar si el proyecto es o no económicamente factible permitiendo elaborar planes de desarrollo. Piense, por ejemplo, en el caso de un proyecto para construir una represa con ciertas especificaciones, o una nueva nave espacial, o un rascacielos, etc.

Ayudan a prevenir circunstancias de emergencia, verificando las medidas más correctas antes de que la situación se presente. Suponga el estudio de un nuevo aparato de buceo; el modelo permitiría determinar reacciones del aparato sometido a ciertos niveles de presión y velocidad. El poder prevenir ciertas reacciones de crisis hará posible realizar cambios en el proyecto para su mejor desarrollo. (ver Figura 3.4)



Figura 3.4: Ejemplo de Ayuda al Desarrollo

3.3.5 Una Ayuda a la Experimentación

Experimentar en este caso significa “investigar, analizar, probar, estudiar, escrutinar, sopesar con un modelo en vez del sistema real”.

La simulación entonces permite probar y evaluar nuevos sistemas o cambios propuestos en los sistemas actuales, ya que experimentar con el sistema real es demasiado costoso, poco práctico o ambos. En general, al experimentar en un modelo de un sistema complejo, podemos aprender más acerca de su interacción interna y de sus elementos, de lo que aprenderíamos por medio de la manipulación del sistema real. Esto se debe al control de la estructura organizacional del modelo, su mensurabilidad, y por la facilidad de variación de sus parámetros.

Las funciones del modelo antes descrito nos permiten hacer deducciones acerca del sistema, dándonos la ventaja de trabajar con él:

- Sin necesidad de reproducir el sistema, debido a que sólo se tiene que plantear como un modelo.

- Sin perturbar el sistema; podemos experimentar con el modelo para analizar su comportamiento, al variar sus parámetros³.
- Sin destruirlo, si estamos simulando el sistema para determinar los límites de resistencia del mismo.

3.4 Clasificación de los Modelos

Los modelos pueden clasificarse de diversas formas. Desafortunadamente ninguna es completamente satisfactoria, aunque cada una sirve a un propósito particular. Es necesario conocer las diferentes clasificaciones porque sólo así sabremos qué tecnologías debemos utilizar. Algunos de los esquemas de clasificación son los siguientes:

1. Clasificación de acuerdo al enfoque del modelo.
2. Clasificación de acuerdo a los cambios o variaciones del modelo.
3. Icónicos versus Físicos versus Análogos.
4. Determinísticos versus Estocásticos.

3.4.1 Clasificación de Acuerdo al Enfoque del Modelo

Esta clasificación se puede esquematizar de la siguiente forma:

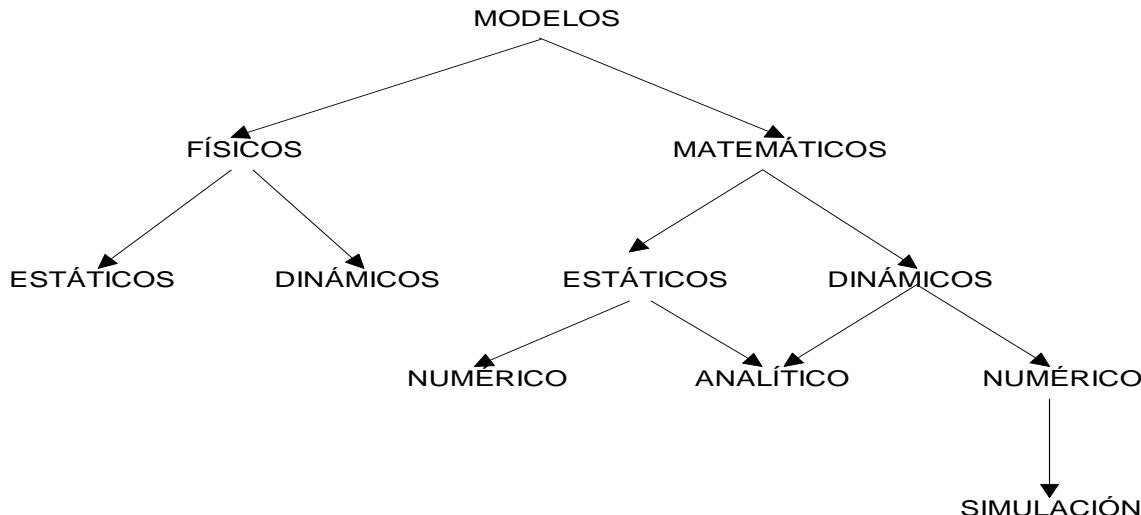


Figura 3.5: Tipos de Modelos

El enfoque de este libro tiene que ver primordialmente con modelos matemáticos, dinámicos, y numéricos.

³ Variables independientes.

3.4.1.1 Modelos Físicos

Los modelos físicos se asemejan al sistema en estudio, y se pueden basar en alguna analogía entre sistemas. En otras palabras, son réplicas o quasi-réplicas del sistema real. Por ejemplo, los sistemas eléctricos, mecánicos o hidráulicos, en donde las actividades del sistema son reflejadas por las leyes físicas que dirigen el modelo, y los atributos del sistema se representan por medidas tales como el voltaje, posición de un eje o motor eléctrico.

Estos modelos, pueden basarse en representaciones a escalas, como las maquetas de edificios, maqueta de una red de tuberías, etc. Este tipo de representaciones se utilizan usualmente para ver la proporción de espacio que hay entre sus elementos. Los modelos físicos se usan tanto en demostraciones como en la experimentación directa. Se les conoce también con el nombre de *íconicos* (ver Figura 3.6).

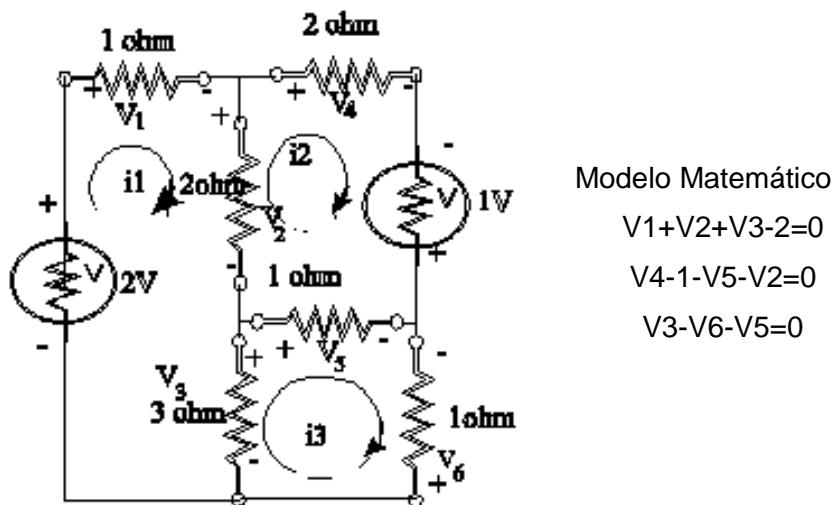


Figura 3.6: Modelo a Escala

3.4.1.2 Modelos Matemáticos

Se utiliza notación simbólica y ecuaciones para representar los modelos. Los atributos del sistema son representados por funciones⁴ matemáticas que indican las relaciones entre las variables. Son más generales que otros tipos de modelos. Existe peligro en su uso por su carácter de idealización abstracta de un problema, porque requieren de suposiciones y de simplificaciones para ser resueltos. Debe tenerse especial cuidado para asegurar que el modelo sea una representación válida del sistema (ver Figura 3.7).

⁴ Fórmulas



Modelo Matemático

$$V_1 + V_2 + V_3 - 2 = 0$$

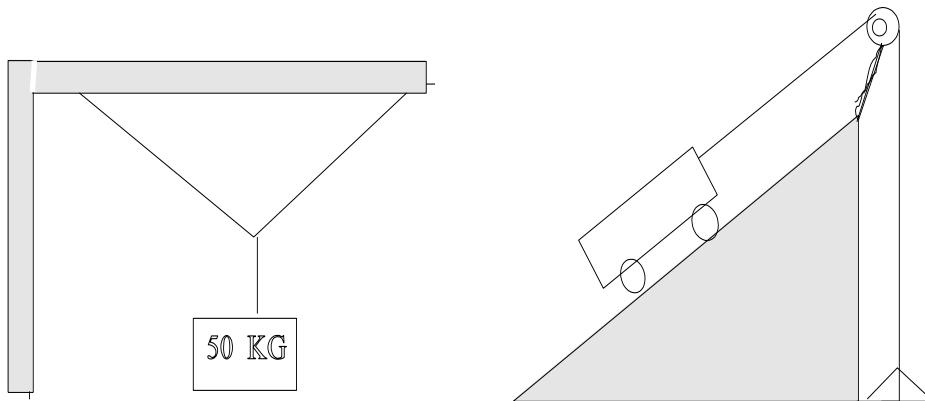
$$V_4 - 1 - V_5 - V_2 = 0$$

$$V_3 - V_6 - V_5 = 0$$

Figura 3.7: Ejemplo de Modelo Matemático

3.4.1.3 Modelos Estáticos

Los modelos estáticos son aquellos que muestran los atributos del sistema solamente cuando este se encuentra en equilibrio. En la investigación de operaciones, con raras excepciones, la mayoría de los trabajos en el área de programación lineal, se han basados en modelos estáticos (ver Figura 3.8).

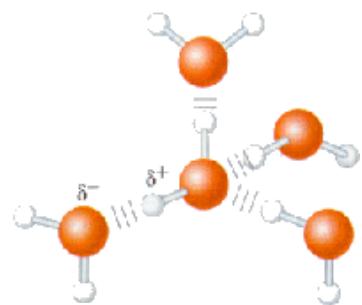
Figura 3.8: Ejemplo de un Modelo Estático:
Dos Cuerpos en Puntos de Equilibrios

3.4.1.4 Modelos Dinámicos

Siguen los cambios del sistema a través del tiempo como resultado de sus actividades. Pueden ser con modelos físicos o matemáticos.

3.4.1.5 Modelos Físicos Estáticos

Son semejantes al sistema real y lo representan en un punto de equilibrio. Los más conocidos son los modelos a escala usados para representar barcos y aviones. En el área científica se utiliza este tipo de modelo para representar el átomo y un modelo de esta naturaleza jugó un papel importante para descifrar la molécula del ADN. Son usados porque resultan más convenientes que tener que producir dibujos complicados en tres dimensiones. Es posible que tiendan a desaparecer ciertos niveles por la facilidad de representación tridimensional de las computadoras.



3.4.1.6 Modelos Físicos Dinámicos

Son aquellos que cuentan con alguna analogía entre el sistema estudiado y algún otro sistema de diferente naturaleza. Esta analogía depende de la semejanza de las fuerzas que gobiernan el comportamiento de los sistemas, y pueden entrar en la categoría de modelos análogos. Un ejemplo de modelo físico-dinámico es un modelo a escala de un avión probado en un túnel aerodinámico para examinar su estabilidad. La siguiente figura muestra un modelo físico con la capacidad de movimiento que permite analizar sus cambios en el tiempo (ver Figura 3.9).

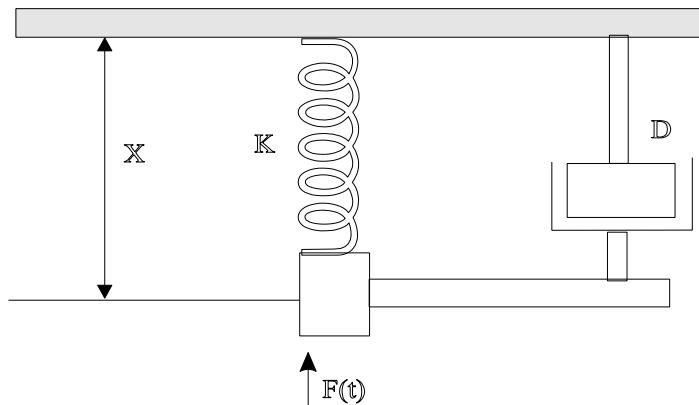


Figura 3.9: Ejemplo de Modelo Físico Dinámico

El diagrama representa un sistema de amortiguamiento cuyo comportamiento varía de acuerdo a la ecuación:

$$Mx'' + Dx' + Kx = KF(t)$$

donde X = distancia

M = masa

K = constante del resorte

D = factor de amortiguamiento

El modelo bien puede estar representando un sistema de amortiguamiento para un determinado automóvil, para mejorar el diseño, aunque sus estructuras no sean idénticas.

3.4.1.7 Modelos Matemáticos Estáticos

Estos modelos indican las relaciones entre los atributos del sistema cuando este se encuentra en equilibrio. Si el punto de equilibrio cambia, por la alteración de cualquiera de sus atributos, el modelo habilita los nuevos valores para todos los atributos en estudio; pero no muestra la forma en la cual ellos cambian a sus nuevos valores. Por ejemplo, en mercadeo, un artículo tiene un equilibrio entre la oferta y la demanda y ambos factores dependen del precio. Un modelo de mercadeo sencillo muestra el precio en el cual ocurre el equilibrio.

A continuación, en la figura 3.10, mostramos una gráfica de la relación entre la oferta (Q) y el precio (P).

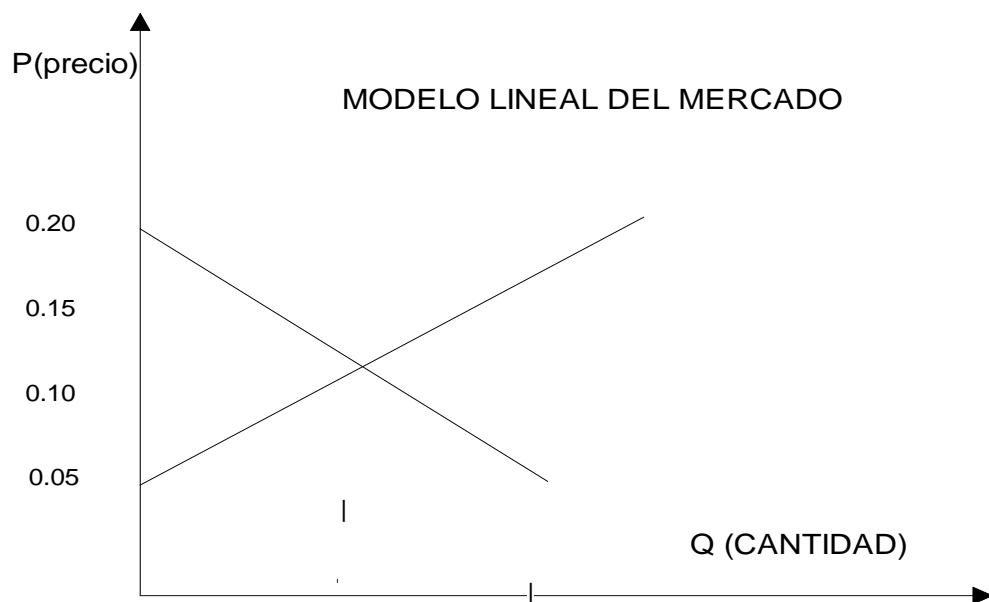


Figura 3.10: Ejemplo de Modelo Matemático Estático

Se aprecia que la demanda del producto será baja cuando el precio sea alto y aumentará al caer el precio.

3.4.1.8 Modelos Matemáticos Dinámicos

Muestran los cambios de los atributos de un sistema como una función del tiempo. La derivación puede hacerse como una solución analítica o con una computación numérica,

dependiendo de la complejidad del modelo. La ecuación que describe la suspensión de un auto es un ejemplo de este tipo de modelo, y esta puede resolverse analíticamente.

$$X'' + 2zW X' + W^2 X = W^2 F(t)$$

Donde $2zW = D/M$

$$W^2 = K M$$

Expresado en la forma anterior se puede obtener una solución en términos de la variable wt

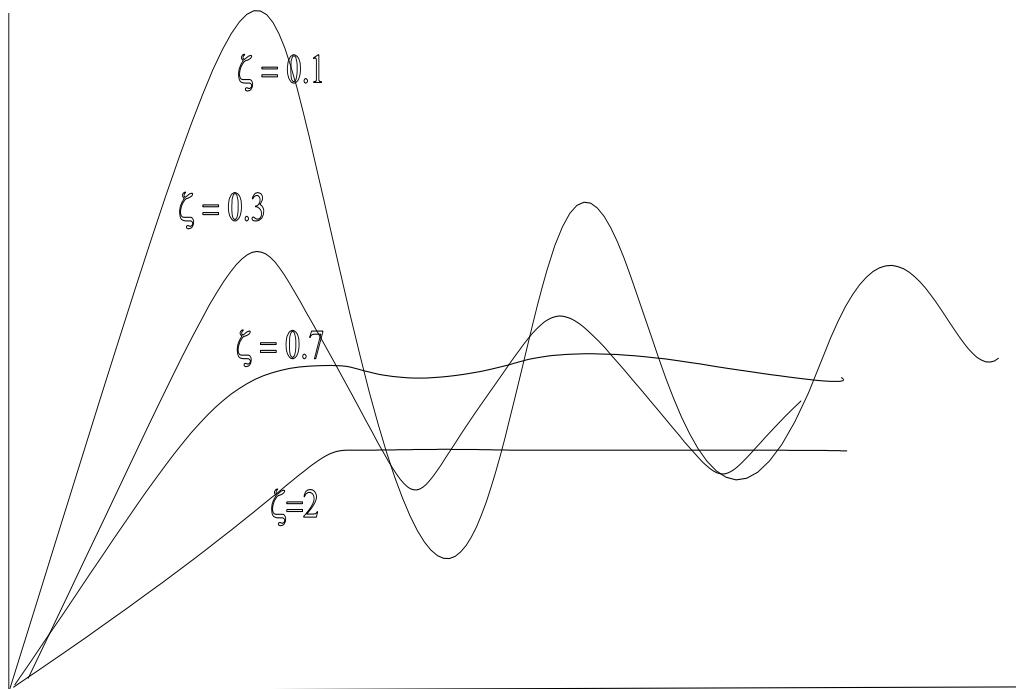


Figura 3.11: Ejemplo de Modelo Matemático Dinámico

La figura 3.11 muestra como x varía en respuesta a una fuerza constante aplicada al tiempo $t=0$; como ocurriría si una carga se monta repentinamente sobre el automóvil. En la gráfica se muestra la solución para algunos valores de z . Cuando z es menor que 1, se trata de un fenómeno oscilatorio.

El factor z es llamado tasa de amortiguamiento. Si el movimiento es oscilatorio, la frecuencia de oscilación se determina por la fórmula: $W = 2\pi f$, donde f es el número de ciclos por segundos.

3.4.1.9 Técnicas Numéricas usadas para resolver Modelos Matemáticos

Involucra la aplicación de procedimientos computacionales⁵ y tablas matemáticas para resolver las ecuaciones.

Para $n = 0, 1, 2, 3, \dots$, hasta donde satisfaga, hacer

Calcular $X_{n+1} = X_n - f(X_n) / f'(X_n)$

Algoritmo del método de Newton dada $f(X_n)$ continuamente diferenciable y un punto X_0

3.4.1.10 Técnicas Analíticas usadas en Modelos Matemáticos

Involucra el razonamiento deductivo de las teorías matemáticas para resolver el modelo. Son convenientes cuando se inicia el estudio del sistema. Estas técnicas producen soluciones manejables y sólo pueden resolver ciertas formas de ecuaciones. Por ejemplo, las ecuaciones diferenciales lineales.

3.4.2 Clasificación de acuerdo a los Cambios o Variaciones del Modelo

Esta clasificación tiene que ver con los cambios *continuos* o *discretos* que ocurren en el modelo.

Las variables independientes son aquellas que pueden tomar cualquier valor porque no están sujetas a ninguna regla o dependencia. Y las variables dependientes son las que están relacionadas con las variables independientes matemáticamente, tomando valores dependiendo de esta⁶. En muchas simulaciones, el tiempo es la variable independiente y las restantes que dependen de él, son las variables dependientes. El comportamiento de las variables dependientes está determinado por el tipo de simulación ya sea discreto o continuo. En realidad, es posible modelar el mismo sistema con cambios continuos o discretos en el modelo. Todo depende del enfoque y objetivo que se le esté dando a la simulación.

3.4.2.1 Simulación Discreta: Modelo Discreto

En la simulación discreta, el sistema simulado se observa únicamente en puntos seleccionados en el tiempo, o sea, que es cuando las variables dependientes cambian discretamente, en puntos específicos del tiempo. Un ejemplo común de una simulación discreta es un sistema de líneas de espera en el cual se unen clientes a una línea de espera o se les ofrece servicio y abandonan la instalación de servicio después de ser

⁵ Hojas electrónicas como MS Excel.

⁶ Rubén Darío, Gran Diccionario Encyclopédico, p. 1288.

atendidos. Supongamos que se está simulando el cajero de un banco⁷. Las variables dependientes serían el estado del cajero y el número de clientes esperando. Los eventos en el tiempo corresponderían a los tiempos de llegada de clientes y también su salida del sistema, después de ser atendidos por el cajero. Esto ocurre en puntos específicos del tiempo.

En general, los valores de las variables dependientes en los modelos discretos no cambian entre eventos en el tiempo. Es decir, mantienen un valor constante a lo largo de cada subintervalo de tiempo. La siguiente figura muestra la respuesta de una variable dependiente en una simulación discreta.

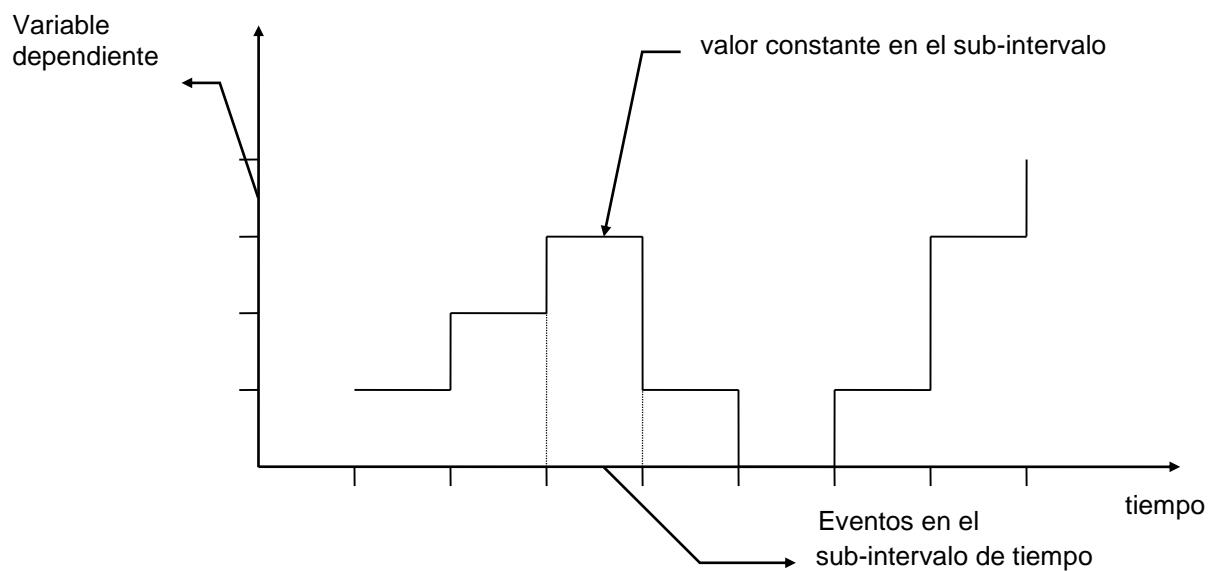


Figura 3.12: Datos Discretos

En la figura 3.12 se muestran medidas de las respuestas de un Simulador de Eventos Discretos. Se puede observar que durante cada intervalo de tiempo se tiene una cantidad dada de clientes.

Es importante mencionar que en algunos casos se usan modelos con una representación continua, cuando en realidad se están simulando sistemas cuyas variables son discretas. Como en el modelo las variables se desarrollan como continuas, diríamos entonces que estamos realizando una simulación continua. Este tipo de experimentación se realiza cuando las entidades en el sistema se consideran en una forma global, antes que como una entidad individual. Por ejemplo, al modelar la población de animales en un lago, el modelador puede usar una representación continua, aunque ésta cambia realmente en una forma discreta.

Un ejemplo más detallado de simulación discreta se presenta en el punto 3.7 de este capítulo.

⁷ Taha Hamdy A., “Investigación de Operaciones”, pp. 653-712.

3.4.2.2 Simulación Continua: Modelo Continuo

Hablamos de simulación continua cuando las variables dependientes del modelo cambian continuamente durante el tiempo simulado. Por lo general, los modelos de simulación continuos involucran una o más ecuaciones diferenciales que indican la relación entre las variables con respecto al tiempo.

Un modelo continuo puede representarse como continuo o discreto en el tiempo. Estos dos casos se presentarán o no, de acuerdo a si se les permite o no, a las variables dependientes, presentarse en todos los puntos del tiempo o solamente en puntos específicos de la simulación. Ejemplos de situaciones donde resulta apropiada una representación continua son: El modelo de la concentración de la reacción de un proceso químico y la posición y velocidad de un avión. Otro ejemplo de simulación continua puede ser la variación de la temperatura estacional, para ser más específica una representación de las temperaturas medias diarias para Dayton, Ohio de enero de 1978 a septiembre de 1979⁸. Se puede representar como continuo o discreto en el tiempo (ver figuras 3.13 y 3.14).

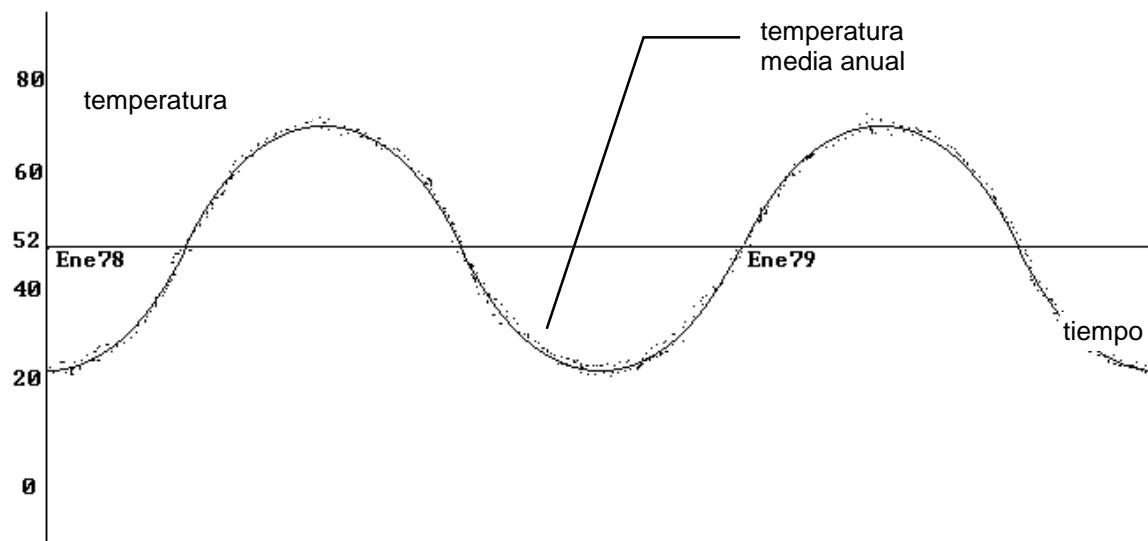


Figura 3.13: Datos Continuos

Se puede observar en estas figuras las mediciones de la respuesta de un simulador de eventos continuos.

⁸ Rice Bernard, "Trigonometría Plana", pp. 170-171.

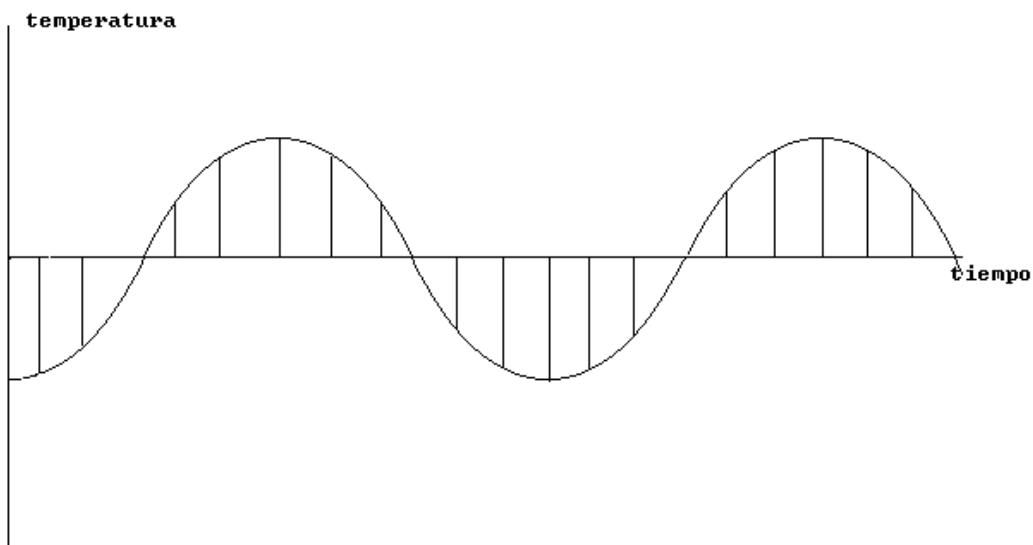


Figura 3.14: Datos Continuos

Mediciones de la respuesta de un simulador continuo usando intervalos de tiempo discreto.

3.4.2.3 Simulación Combinada: Modelo Combinado

Se dice que una simulación es combinada cuando algunas variables dependientes de un modelo pueden cambiar discretamente y otras continuamente.

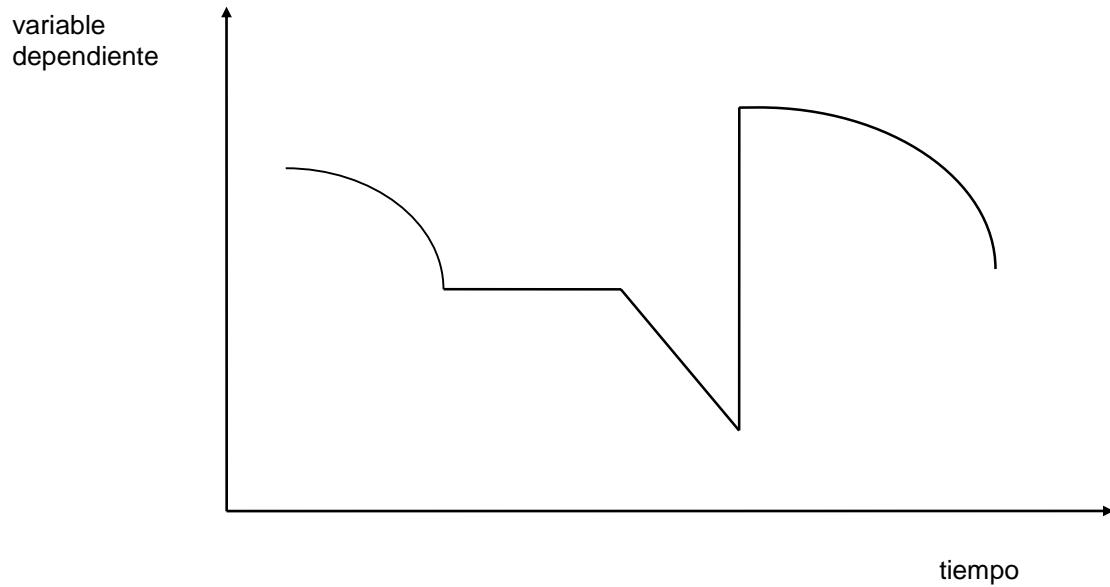


Figura 3.15: Medición de la Respuesta de un Simulador Combinado

También puede variar continuamente con puntos discretos en el tiempo. El tiempo puede considerarse discreto. La interacción que tiene lugar entre las variables con cambios discretos o continuos es lo más importante en este tipo de combinación.

3.4.3 Otra Clasificación

Una tercera clasificación agrupa a los **modelos como icónicos versus simbólicos versus análogos**. Los tres tipos fueron descritos previamente en la primera clasificación, ya que los modelos icónicos son los físicos, los simbólicos son los matemáticos y algunos modelos análogos son también físicos. Sin embargo, hemos querido ampliar un poco más en esta sección la explicación de modelos análogos.

Modelos análogos son aquellos en los cuales una propiedad de un objeto real se representa por una propiedad substituida, la cual presenta un comportamiento similar. El problema se resuelve en el estado substituto y se trasladan las respuestas a las propiedades originales.

Por ejemplo, en una computadora análoga, el voltaje a través de una red (malla) puede representar el flujo de elementos, tales como: artículos, datos, productos, etc. a través de un sistema. Otro ejemplo de modelo análogo muy usado son las gráficas. En ellas, las distancias muestran propiedades tales como el tiempo, edad, etc., presentando la relación entre diferentes cantidades. Una de sus ventajas es que ayudan a predecir los cambios entre estas cantidades cuando alguna de ellas varía.

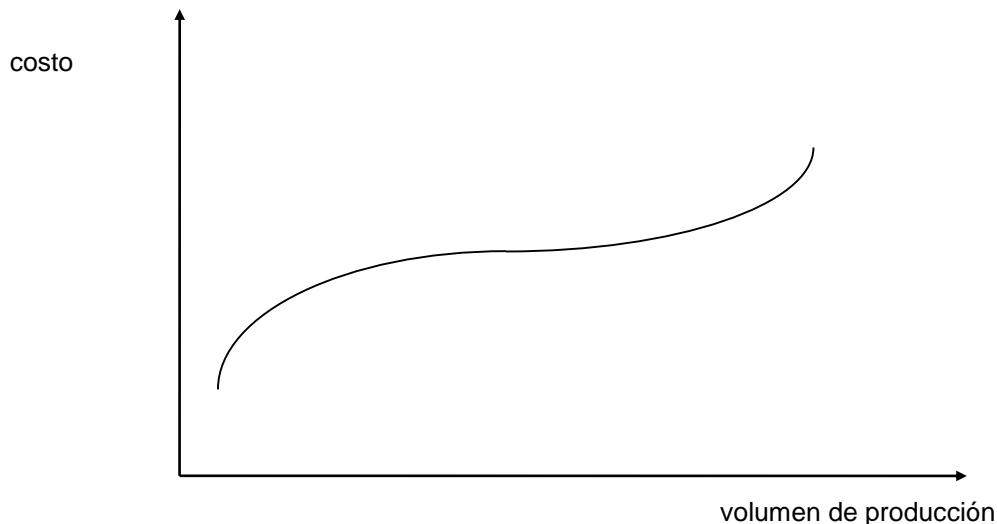


Figura 3.16: Curva del Costo de Producción

En la figura 3.16 se aprecia la relación entre el costo de producción y el volumen. Podemos predecir lo que ocurrirá con el costo cuando el volumen aumenta o disminuye. Estos modelos permiten resolver problemas relativamente simples.

En resumen, podemos decir que es difícil establecer el modelo de un sistema complejo. Sin embargo, el investigador puede recurrir a una combinación de más de uno de los

tipos de modelos discutidos, para el estudio de un mismo sistema. Esta combinación puede variar ampliamente en complejidad y configuración. “Para un mismo sistema se pueden realizar muchos y muy diversos estudios a partir de los modelos que de él se construyan”.

Por lo general, los modelos simples conducen a modelos más complejos de acuerdo a cómo el investigador continúe analizando y entendiendo más profundamente el problema.

Cuando se desarrolla un modelo de simulación, el modelador necesita establecer un marco de trabajo, para describir el sistema que va a modelar. Este marco de trabajo no es más que la perspectiva dentro de la cual se perciben y describen las relaciones funcionales del sistema. Si el modelador utiliza un lenguaje de simulación la perspectiva mencionada estará implícita en él.

3.5 Esquema de la Implementación de un Modelo

El diagrama que se muestra a continuación (ver figura 3.17) ilustra la posición que ocupa un modelo con respecto al sistema que representa y a su implementación. Es importante resaltar que:

- El modelo se encuentra en un punto intermedio entre el sistema y su implementación.
- El modelo toma del sistema elementos tales como un propósito, sus entidades y atributos más importantes y pertinentes, sus límites, alternativas, etc. La evaluación rigurosa del modelo determina su validez y calidad.
- La implementación del modelo se realiza a partir de las evaluaciones que resultan de la simulación.

La solución de un problema a través de la simulación implica el seguimiento de una serie de pasos, los cuales son explicados en la siguiente sección.

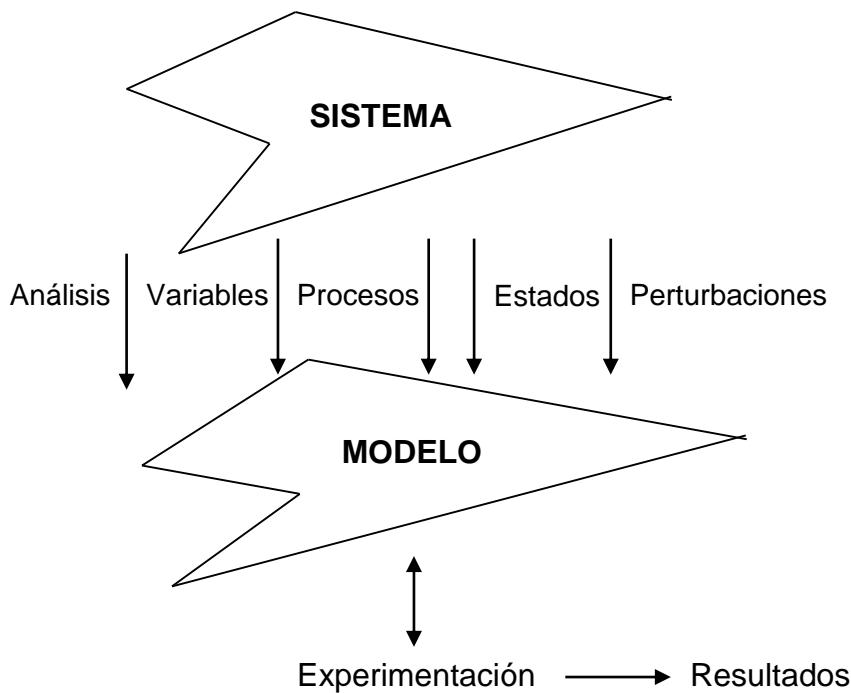


Figura 3.17: Esquema de la Implementación de un Modelo

3.6 Etapas en la Solución del Problema

El éxito de una simulación consiste en iniciar un modelo e ir mejorándolo progresivamente al ir adecuando los requerimientos para la solución del problema. Dentro de este contexto podemos identificar las siguientes etapas:

3.6.1 Definición del Problema

Hay que empezar por saber cuál es el problema; en otras palabras, hay que definirlo. Es necesario, pues, que se identifiquen las condiciones para que el problema exista.

Para realizar la definición del problema primero hay que establecer los objetivos, así como las condiciones de frontera. Si se dan las condiciones necesarias, se inicia un proceso interactivo entre quienes tienen el problema y quienes van a construir el modelo.

Debido a la naturaleza evolutiva de la simulación en que la investigación genera nueva información referente a restricciones, objetivos y alternativas, definir un problema es un proceso continuo que ocurre durante todo el estudio. Podemos describir o definir lo que ocurre dentro de las fronteras de un sistema de varias formas diferentes. La tendencia es hacerlo con muchos detalles. Pero lo cierto es que, si no se seleccionan los elementos y las relaciones por considerar para un propósito específico, sería posible un número infinito de conexiones y combinaciones. Por lo tanto, el modelo debe incluir sólo aquellos aspectos importantes del sistema para el estudio de los objetivos.

La definición del problema se realiza a partir de las vistas o revisiones internas del problema y a través de preguntas adecuadas. Imaginemos que el problema que se nos ha presentado consiste en la construcción de un nuevo puente sobre el Canal de Panamá. Los puntos que ayudan a definir los objetivos son los siguientes:

- **Preguntas que deben contestarse:** Este es un conjunto refinado de preguntas que requieran de secuencias lógicas.

Por ejemplo:

¿Es necesario un nuevo puente sobre el Canal de Panamá?

¿Dónde se debe construir un nuevo puente?

- **Hipótesis que deben verificarse o refutarse:** La formulación del problema se logra aceptando o refutando un conjunto de hipótesis. Por ejemplo:

¿La construcción del nuevo puente acelerará el desarrollo o deteriorará la calidad de vida de sus habitantes?

¿El nuevo puente acortará el tiempo de tránsito de los habitantes que comutan entre sus hogares y la Ciudad de Panamá?

- **Efectos que hay que estimar:** Es necesario estimar los efectos que tienen las actividades exógenas en el sistema bajo estudio. Ejemplo:

¿Cómo afectará al transporte foráneo que llega a la ciudad, el hecho de que el nuevo puente sobre el canal se localice a la altura de Colón?

¿Cuánto perjudicará el nuevo puente al ecosistema?

En este paso la participación activa de los que tienen y entienden el problema es una condición necesaria.

3.6.2 Creación del Modelo

Una vez definido el problema se inicia la etapa de creación del modelo. El modelo consistirá en una representación estática o dinámica del sistema en estudio. Este proceso es realmente un arte. El modelador debe entender la estructura y reglas de operación del sistema. También debe estar dispuesto a extraer la esencia del sistema sin incluir detalles innecesarios.

Un modelo ideal es aquel que resulta fácil de entender y lo suficientemente complejo como para reflejar realmente las características más importantes del sistema. Las decisiones cruciales necesarias para simplificar el modelo son muy importantes, al igual que la cantidad de detalles que se incluirán de acuerdo a los propósitos del estudio. Es importante considerar sólo aquellos elementos que puedan causar diferencias significativas en la toma de decisiones.

Por eso el modelador debe:

- Establecer el propósito para la creación del modelo.
- Determinar si una entidad (elemento u objeto) es lo suficientemente significativo de acuerdo al propósito establecido.
- Determinar cuáles son las características más importantes de cada entidad.
- Establecer las relaciones entre estas entidades.

En primer lugar, el establecimiento de las estructuras del modelo, consiste en determinar:

- Entidades (que son las partes que ejecutan alguna función o proceso).
- Fronteras: (los límites del sistema: qué forma y qué no forma parte del sistema por estudiar).
- Atributos (manifestaciones apreciables o propiedades que describen a las entidades).
- Actividades (proceso que causa cambios en el sistema).
- Otros.

Los aspectos anteriores fueron descritos con mayor detalle en el capítulo 1.

El éxito del modelador depende de la exactitud que logre en cada uno de los puntos antes mencionados. En el ejemplo del sistema de tránsito que involucra el área del canal, podemos identificar algunos elementos:

Entidades	Atributos	Actividades
Vehículos	Peso, número de pasajeros, tipo	Desplazarse de un lado al otro del puente
Señales	Distancia (distribución),	
Puente	Localización geográfica, número de carriles, Longitud	

Aquí se escoge alguna forma de representación, una gráfica, maqueta, conjunto de ecuaciones, etc.

Toda la formulación del problema como la del modelo requieren de una interacción estrecha por parte del personal del proyecto. En su etapa inicial, la construcción, análisis y discusión del modelo requerirá de suposiciones apropiadas por parte del modelador, que a su vez debe estar dispuesto a experimentar su ignorancia potencial del sistema. Por ejemplo, no podemos conocer inicialmente la preferencia de los conductores sobre la utilización del Puente de las Américas o el nuevo puente en estudio, así como cualquier otro factor que no nos imaginemos pero que pueda aparecer durante el desarrollo.

3.6.3 Adquisición⁹ de Datos

El suministro de los datos que va a utilizar el modelo es muy importante y se cuenta con tres posibles fuentes para generarlos:

- Datos históricos y series de tiempo: Recopilaciones de datos realizadas a través del tiempo.
- Opiniones de expertos: La experiencia puede suministrar valiosa información y bastante acertada.
- Estudios de campo: Un estudio del problema en cuestión y los datos se recolectan donde se encuentra el sistema (el capítulo 4 explica esta opción en más detalle).

El modelador debe preocuparse por los datos referentes a las entradas y salidas del sistema estudiado, así como la información acerca de los componentes del sistema y de las interconexiones o relaciones entre ellos. Por lo tanto, se interesa en la recolección de datos cuantitativos y cualitativos, y debe decidir qué datos se necesitan, si son importantes, si los datos que existen son válidos para sus propósitos, y cómo recopilar esta información¹⁰. Involucra la identificación, especificación y colección de datos. Algunos se pueden tener en el acto, mientras que otros pueden involucrar un tiempo y costos considerables para su recolección. Estos datos pueden ser hipotéticos inicialmente o se basan en un análisis preliminar. En muchas ocasiones los valores exactos para uno o más datos de entrada pueden tener pequeños efectos o variaciones en los resultados de la simulación. La sensibilidad en los resultados de la simulación por cambios en los datos de entrada del modelo, pueden evaluarse realizando una serie de corridas de simulación variando los valores de los parámetros de entrada. Puede usarse el modelo para determinar cuál es la mejor forma de asignar el dinero y el tiempo en lo referente a la entrada de datos. Un fracaso común en los estudios de simulación es que se concentra tanto en ésta, que de la simulación se extraen más datos de los necesarios o de los que pueden validarse con los datos disponibles¹¹.

Por ejemplo, para nuestro modelo de simulación del tránsito en un puente nuevo sobre el Canal de Panamá, podríamos usar valores recogidos por la Autoridad Nacional de Tránsito y Transporte Terrestre en períodos normales sobre el Puente de las Américas y en períodos críticos como carnavales, Semana Santa, Navidad, fines de semana largos. Podríamos recoger datos tomándolos nosotros mismos sobre el puente actual y luego agrupándolos.

3.6.4 Codificación del Modelo

En esta etapa preparamos el modelo para los procesos de computación. Una vez desarrollado el modelo y establecidas las estimaciones iniciales de entrada, el siguiente

⁹ Recolección.

¹⁰ Shannon R., "Simulación de Sistemas", p. 39.

¹¹Gordon Geoffrey. Simulación de Sistemas. Pág. 38

paso es traducir el modelo a una forma aceptable por la computadora. En otras palabras, es traducir el modelo a un lenguaje entendible por la computadora.

Dado un modelo matemático, la construcción de un programa de computadora para el modelo es una tarea relativamente bien definida, no implicando que sea fácil de hacer. Es probable que la tarea de formular el modelo se haga a la vez que el traspaso a la computadora y no separadas¹².

Aunque un modelo de simulación puede programarse usando un lenguaje de propósitos generales, hoy en día existen indudables ventajas al usar un lenguaje de simulación; los cuales nos ahorran tiempo de programación y ayuda en la formulación del modelo. Esta ayuda la proveen mediante un conjunto de conceptos que permiten articular la descripción del sistema. Más adelante se presentan los lenguajes de simulación ARENA y GPSS.

3.6.5 Verificación y Validación

Este es un punto donde se requiere de un buen juicio, ya que los simuladores pueden parecer reales, y tanto los modeladores como los usuarios pueden encontrarlos fácilmente creíbles. Lamentablemente, con frecuencia los simuladores ocultan sus suposiciones del observador casual y algunas veces del modelador. Por tanto, si no se llevan a cabo cuidadosa y detalladamente, pueden aceptarse resultados erróneos con consecuencias desastrosas¹³.

La *verificación* consiste en determinar si el modelo realiza las ejecuciones en la computadora como ideó el modelador. *Validar* es determinar si el modelo es una representación razonable y aceptable del sistema. Es recomendable validar las entradas, los elementos del modelo, los subsistemas y puntos de interfaces.

Ambas no deben confundirse. La *verificación* implica la revisión de la lógica usada mientras que la *validación* compara los resultados del modelado con resultados históricos producidos por el sistema real, bajo las mismas condiciones.

En conclusión, se deben responder preguntas como: ¿es posible que el modelo dé respuestas absurdas si se lleva a los parámetros a valores extremos? También, si los resultados del modelo parecen razonables. Esto se puede resolver encontrando gente que intervengan directamente en el sistema real y solicitarles que comparan los resultados del simulador con las salidas reales. Para hacer esto un poco científico, se suministran datos reales y datos simulados a gente experimentada para ver si identifican las diferencias.

Por ejemplo, cuando se nos pide un programa sobre inventario, este se prueba antes de entregarlo para confirmar que corre en la forma adecuada. Al verificar, estamos

¹²Gordon Geoffrey, Simulación de Sistemas, Pág. 38

¹³ Shannon Robert, "Simulación de Sistemas", p. 41.

evaluando si ejecuta todas las rutas lógicas posibles y su forma como se describió. Al comparar los resultados, estamos validando.

3.6.6 Planeación de Técnicas y Estrategias

Consiste en desarrollar procesos que ayudan a establecer las condiciones experimentales o ambiente preciso, para las corridas de simulación. Se prepara un diseño experimental para explicar cualquiera de las relaciones entre la respuesta de simulación y las variables controlables, o para encontrar la combinación de valores para las variables controlables, con las cuales minimizar o maximizar respuestas de simulación.

En otras palabras, la planeación estratégica es la recopilación de información original que proporcionará el suficiente conocimiento acerca de los fenómenos del sistema, para permitir que se obtengan inferencias válidas acerca de su desempeño. Esto pudiéndose lograr ya sea encontrando la combinación de valores de parámetros que optimizarán a la variable de respuesta, o explicando la relación entre la variable de respuesta y los factores controlables dentro del sistema. Y la planeación técnica controla los aspectos de eficiencia a través del control de cómo se llevará cada una de las corridas del simulador que se especifican en la planeación estratégica.

Un factor que debe considerarse es el costo que se incurre al realizar una cantidad grande de experimentos, por tal razón se debe planear cuántos de ellos deben ser corridos y que cumplan con los objetivos propuestos. Hay casos en que este factor no es de mayor importancia, pero aun así debe considerarse la planeación debido a que el estudio podría fallar debido a que el modelador se encuentra abrumado con la excesiva cantidad de resultados de la computadora¹⁴.

En consecuencia, la planeación puede mejorar notablemente la síntesis del nuevo conocimiento y la conjetura de nuevas ideas, y al mismo tiempo, minimizar el esfuerzo, tiempo y costo del experimento.

3.6.7 Experimentación y Análisis de Resultados

En esta fase se ejecuta el modelo y se interpreta los resultados de la corrida. Cuando los resultados de la experimentación se usan para desarrollar suposiciones o pruebas de hipótesis, se emplean los métodos estadísticos. O sea, se pasarán a realizar los experimentos que han sido planeados en el paso anterior.

Se ejecutarán las corridas de simulación e interpretación de los resultados con la intención de responder a un conjunto bien definido de preguntas que se hayan planeado con anterioridad para llegar a tomar hipótesis¹⁵. En otras palabras, la experimentación es el proceso de observación y análisis que proporciona la información que conduce a las soluciones.

¹⁴ Gordon Geoffrey, "Simulación de Sistemas", p. 39.

¹⁵ Gordon Geoffrey, "Simulación de Sistemas", p. 39.

En esta fase empezamos a encontrar los defectos y errores de nuestra planeación y a volver sobre nuestros pasos hasta alcanzar los objetivos originalmente establecidos (ver figura 3.18).

3.6.8 Implementación de los Resultados y Documentación del Modelo

La etapa final de este proceso es la implementación de resultados conjuntamente con la documentación. No se puede considerar que un modelo de simulación está completo hasta que sus resultados son aceptados, entendidos y usados en el proceso de toma de decisiones.

Las actividades en este punto son prolongadas y dependen del grado en el cual el modelador haya ejecutado acertadamente las otras actividades en el desarrollo del proceso de simulación. Si el modelador y el usuario del modelo han trabajado conjuntamente y ambos entienden el modelo y sus salidas, entonces es probable que los resultados del proyecto sean implementados con precisión. De otra forma, si no se han comunicado eficientemente la formulación del modelo y los subsiguientes pasos, sería más problemático la implementación de las recomendaciones a pesar de la elegancia y validez del modelo.

La documentación se relaciona estrechamente con la implementación. La documentación cuidadosa y completa del desarrollo y la operación del modelo pueden incrementar notablemente su vida útil y sus oportunidades para una exitosa implementación. La buena documentación facilita la modificación y asegura que el modelo pueda usarse a pesar de que los servicios que se encargan del desarrollo ya no estén disponibles. Además, el llevar una documentación cuidadosa puede ayudarle al modelador a aprender de sus propios errores y quizás proporcionarle una fuente de subprogramas que pueden volverse a usar en futuros proyectos¹⁶.

Las etapas o fases que se describieron, raramente se desarrollan en la secuencia que se ha indicado, ya que un proyecto de este tipo involucra suposiciones erradas, entradas falsas, las cuales se abandonan para formular nuevamente el objetivo del problema y repetir la evaluación y rediseño del problema. Este es un proceso iterativo como se puede apreciar en el siguiente diagrama de flujo (Ver Figura 3.19 Diagrama de Construcción de un Modelo¹⁷).

¹⁶ Shannon, Robert, "Simulación de Sistemas", p. 45.

¹⁷ Idem, p.24.

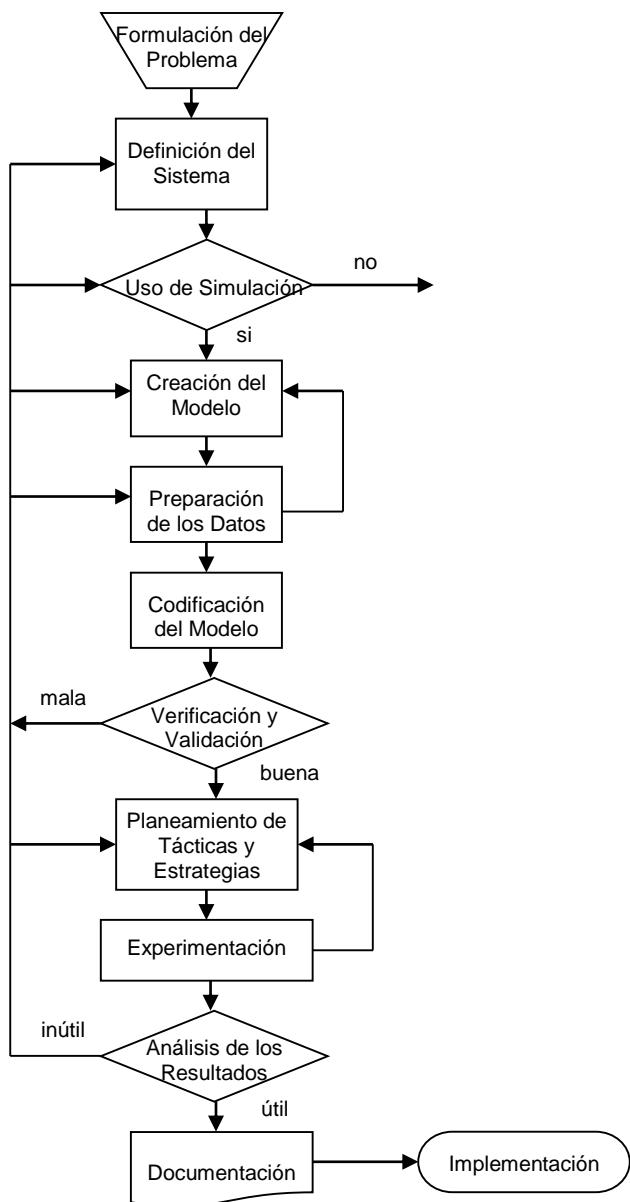


Figura 3.18: Diagrama de Construcción de un Modelo

3.7 El Problema de la Percepción

Percepción se define como el proceso mediante el cual la información que nuestros sentidos captan del medio es interpretada por las funciones básicas de nuestro cerebro. Dicho de otra forma, los datos recibidos por nuestros sentidos no son simplemente registrados en un banco de memoria, sino que les son asignados ciertos atributos predeterminados por nuestros paradigmas, valores, principios y organizados sobre la base de los mismos.

Como vemos las cualidades del modelo descansan en la percepción de quien lo modela puesto que se trata de una abstracción. Si encargamos a varios pintores la ejecución de la representación de un paisaje cualquiera, veremos diferencias apreciables en cada una de las obras, ya sean pintores abstractos o no. Esto es debido al fenómeno de la percepción.

El primer problema de la percepción es que no puede ser comunicada. El estudio de los sistemas mediante los modelos se hace cada vez menos preciso en la medida en que el sistema sea más complejo y que el estudio involucre una mayor cantidad de eslabones en la cadena de investigadores, quienes estarán confinados a emular el uno del otro la percepción que tengan del sistema, e incluso del modelo.

Uno de los objetivos de utilizar un modelo para el estudio de un sistema es poder conocer a fondo las relaciones entre sus elementos. Nace así la discrepancia de cuál será la relación más importante en el estudio para dedicarle mayor tiempo. En sistemas poco complejos esto no parece tener sentido. En el sistema de una Intersección de una Avenida con una Calle (ver capítulo 1) la relación tiempo de espera / cantidad-de-autos-por-el-carril es una, sino la de mayor importancia. Pero en sistemas sociales, las relaciones políticas no requieren tanto estudio como las de salud, en la opinión de un médico, pero sí para un diputado.

Tanto es así, que, dependiendo del modelador, algunas relaciones realmente importantes podrían ser pasadas por alto.

Una serie de experimentos demuestran que “la percepción ya sea en sentido positivo o negativo es influída por factores emocionales. Tendemos a ver aquello para lo que estamos preparados o dispuestos y a no ver lo que nos perturba”.

Para evitar los errores que pueda causar la percepción, es necesario establecer métodos precisos y sistemáticos, acorde con el tipo de sistema que nos lleven por una senda segura de modelado.

3.8 Ejemplo 1: Simulación de Eventos Discretos

Considérese una empresa de servicio con un solo servidor, por ejemplo, un operador de una barbería o una ventanilla donde se suministra información en un aeropuerto, para el cual quisiéramos estimar el promedio de clientes que llegan diariamente. El tiempo que esperara un cliente comienza cuando entra a la cola hasta que empieza a ser servido. Para estimar el tiempo invertido por cada cliente, se establecen variables discretas. Algunas:

- El servidor, ocupado u ocioso,
- El número de clientes esperando en la cola para ser servido, y
- El tiempo de llegada de cada persona en la cola.

El estado del servidor se necesita para determinar, si el cliente que llega será atendido inmediatamente o debe agregarse al final de la cola. Cuando el servidor termina de atender a un cliente, seguirá con el siguiente cliente de la cola, si no hay más clientes el servidor estará ocioso.

Hay dos tipos de eventos para este sistema, llamados:

- Llegadas de clientes (o tiempos entre arribo) y
- Tiempos de servicio de los clientes.

Una llegada de un cliente es un evento porque provoca que el servidor cambie de ocioso a ocupado, o que el número de clientes en la cola aumente en 1.

Correspondientemente, una salida del servidor es un evento, ya que provoca que el estado del servidor cambie a ocioso o el número de la cola disminuye en 1.

Para la solución de este problema necesitamos las siguientes variables:

T_i	= Tiempo de arribo de i -ésimo cliente (se comienza con $t_0= 0$).
A_i	= $T_i - T_{i-1}$ = tiempo entre el arribo del i -ésimo cliente y el (i -ésimo – 1) cliente.
S_i	= Tiempo que el servidor actualmente gasta sirviendo al i -ésimo cliente (excluyendo la demora en la cola del cliente).
D_i	= Demora en la cola del i -ésimo cliente.
C_i	= $D_i + S_i$ = Tiempo que el i -ésimo cliente demora en total (en cola y para ser servido e irse).
S_i	= Tiempo de ocurrencia del i -ésimo evento de cualquier tipo (excluyendo el valor $S_0 = 0$).

Cada una de las colas será definida generalmente por una variable aleatoria. Asumiendo la distribución de probabilidad de los tiempos de intervalos $A_1, A_2\dots$ y la del tiempo de servicio $S_1, S_2\dots$ son conocidas y ambas denotadas por $F(A)$, y $F(S)$, respectivamente. (En general, $F(A)$ y $F(S)$ estarán determinadas por la recolección de datos del sistema de interés y la distribución apropiadas de estos datos¹⁸). En el tiempo $S_0=0$ el estado del servidor es “desocupado”, y el tiempo del primer arribo t_1 , está determinado por la generación de A_1 de $F(A)$ (técnicas para generar valores aleatorios de una distribución específica) y adicionando esto a 0. El reloj de la simulación es entonces incrementado desde S_0 al tiempo del próximo evento, $S_1= T_1$. (Observe la figura 3.19, donde la flecha curva representa el incremento del reloj de simulación). Desde la llegada del cliente en el tiempo T_1 , encontrando al servidor vacío, él es inmediatamente servido y el tiempo de demora en la cola $D_1=0$ y el estado del servidor cambia de ocioso a ocupado. El tiempo, C_1 , cuando el cliente ha llegado hasta que el servicio es completado puede ser computado por la generación S_1 de $F(S)$ y adicionando este a T_1 . Finalmente, el tiempo del segundo entre arribo, T_2 , es calculado como $T_2 = T_1+A_2$, donde A_2 es generado de $F(A)$, si $T_2 < C_1$, como se descubre en la figura 2.20, el reloj de la simulación es

¹⁸ Los capítulos 3 y 4 presentan los conceptos necesarios para determinar las distribuciones de probabilidad del tiempo de entrearribo y del tiempo de servicio.

incrementada de S_1 al tiempo del siguiente evento, $S_2=T_2$. Si C_1 fue menor que T_2 , el reloj será incrementado de S_1 a C_1 . Después el cliente llega en un tiempo T_2 , encontrado al servidor ocupado, el número de clientes en la cola es incrementado de 0 a 1 y el tiempo de llegada de este cliente es almacenado; sin embargo, el tiempo de servicio S_2 no es generado. También, el tiempo del tercer arribo, T_3 , es computado como $T_3 = T_2+A_3$. Si $C_1 < T_3$, como se describe en la figura 2.20, el reloj de simulación es incrementado de S_2 al tiempo de próximo evento, $S_3=C_1$, donde el servicio del cliente es completado, el cliente en la cola empieza a ser servido y su demora en la cola y el tiempo para completar el servicio son computado como $D_2=C_1-T_2$ y $C_2=C_1+S_2$ (S_2 es generado ahora de F_2), y el número de clientes en la cola es decrementado de 1 a 0. Si $T_3 < C_2$, el reloj de simulación es avanzado de S_3 al tiempo del próximo evento, $S_4=T_3$, etc. La simulación puede ser eventualmente terminada cuando se llega al número de clientes, cuya demora tiene que ser observada al alcanzar su valor especificado.

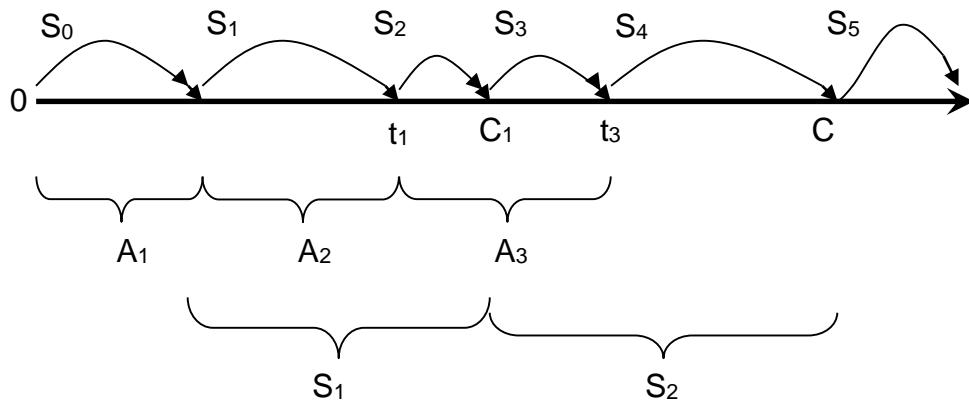


Figura 3.19: Eventos a través del tiempo

3.9 Ejemplo 2: Simulación Continua

Consideremos un modelo continuo, la competencia entre dos poblaciones. Modelo biológico, el cual es llamado modelo “depredador y presa”. Un ambiente consistente en las dos poblaciones, las cuales interactúan entre sí. Las presas son pasivas, pero los depredadores dependen de ella como fuente de alimento. Por ejemplo, el depredador puede ser un león y las presas pueden ser las gacelas de las que se alimenta. Sean cada una $x(t)$ y $y(t)$ respectivamente, el número de individuos de cada población en un tiempo dado. Supóngase que hay un amplio suministro de presas y la ausencia de depredadores, que la tasa de crecimientos es $r_x(t)$, para $r > 0$. Podemos pensar en r como el número de nacimientos menos el número de muertes naturales (rata de crecimiento = rata de nacimientos – rata de defunciones).

De la interacción entre depredador y presa, es razonable suponer que el promedio de muertes de las presas debido a la interrelación es proporcional al producto de las dos poblaciones, dadas por $x(t)$ $y(t)$, por lo que el porcentaje total de cambios de la población, dx/dy , está dada por:

$$\frac{dx}{dy} = rx(t) - ax(t) y(t) \quad (1)$$

donde a es una constante positiva de proporcionalidad.

Como el depredador depende de la presa para su existencia, el porcentaje de cambio del depredador en ausencia de presa es $-sy(t)$, para toda $s>0$. Además, la interacción entre las dos poblaciones causa que la población de depredadores se decremente en un porcentaje el cual es proporcional a $x(t)y(t)$. Debido a esto, el porcentaje total de cambio de la población de depredadores, dx/dy , es

$$\frac{dx}{dy} = -sy(t) + bx(t)y(t) \quad (2)$$

donde b es una constante positiva.

Dadas las condiciones iniciales $x(0)>0$ y $y(0) > 0$, la solución del modelo está dada por las ecuaciones 1 y 2, las cuales tienen que tener la propiedad $x(t)>0$ y $y(t)>0$, para toda $t>0$. Por lo tanto, la población de presa nunca será completamente extinguida por los depredadores. La solución $\{x(t), y(t)\}$ es en sí una función periódica del tiempo. Sea T el período, dado $T>0$, tal que $x(t+nT) =x(t)$ y $y(t + nt) =y(t)$ para toda $n= 1,2,\dots$

Estos resultados no serán inesperados. Cuando la población de depredadores incrementa, la población de presas decremente. Esto causa el decrecimiento de la tasa de crecimiento de los depredadores que a su vez causa que el número de presas incremente.

Considere los valores particulares $r = 0.001$, $a = 2E-6$, $s = 0.01$, $b = 10E-6$ y los valores iniciales del tamaño de las poblaciones $x(0)=12000$ y $y(0) = 600$. La figura 3.20 muestra la solución numérica para este Ejemplo:

Tamaño de la población (En Cientos)

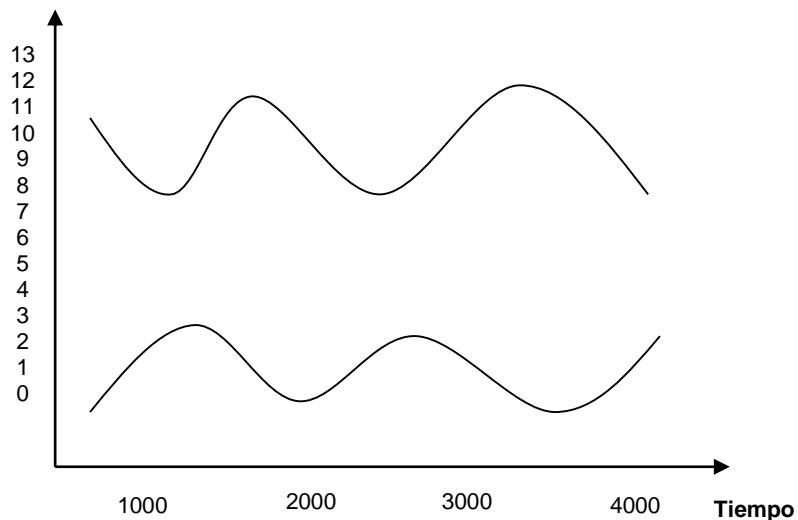


Figura 3.20: Solución Numérica de Depredador y Presa

3.10 Ejemplo 3: Simulación Combinada

Barcos con tanques llenos de aceite crudo llegan a desembarcar a un puerto para aprovisionar un depósito que alimenta una refinería a través de un oleoducto. Una descarga del barco envía aceite al tanque de almacenaje especificada por una tasa constante. Los barcos que llegan cuando el puerto está ocupado se unen a la cola. El depósito suple de aceite a la refinería en diferentes especificaciones de cantidad. El puerto está abierto de 6 a.m. a 12 p.m., por consideraciones de seguridad. La descarga de los barcos cesa cuando el puerto está cerrado.

Los eventos discretos para este modelo simplificado son las llegadas de los barcos tanques para descargar, el cierre del puerto a las 12 p.m. y la apertura a las 6 a.m. Los niveles de aceite descargado de los barcos y del depósito está dada por variables continuas, cuyas variaciones son dadas por diferentes ecuaciones. La descarga de los barcos es considerada como completa cuando el nivel de aceite de los barcos es menor del 5% de su capacidad, pero la descarga de los barcos puede ser detenida temporalmente si el nivel del tanque de almacenaje alcanza su capacidad máxima. La descarga puede ser reasumida cuando el nivel del tanque disminuye a un 80% de su capacidad. Si el nivel de aceite en el tanque cae debajo de los 5000 barriles, la refinería tiene que ser cerrada temporalmente. Con el propósito de evitar los frecuentes cierres y aperturas de la refinería, los barcos no reiniciarán el abastecimiento hasta que este contenga 50,000 barriles. Cada uno de los cinco eventos concernientes a los niveles de aceites, por ejemplo, el nivel de aceite de los barcos cae debajo del 5% de su capacidad, llamado estado de eventos. Puede ocurrir un estado de evento no planeado, pero estos ocurren cuando un estado de las variables continuas cruza su límite establecido.

3.11 Ejemplo 4: Simulación de un Servicio de Respuesta para Consultas a un Centro de Información

Este ejemplo se seleccionó con el propósito de proporcionar al lector una visión global de los tipos de asuntos que el modelador debe enfrentar y resolver. Particularmente nótese la discusión de la *recolección de datos* y la *planeación táctica*.

El centro de información de partes del programa APOLO es un centro de información para recabar, almacenar y divulgar la información sobre las partes y materiales de dicho programa. Específicamente, el centro se interesa sobre las partes y materiales, tales como identificación de partes, información de requerimientos, datos de calidad, datos de prueba, datos de fallas, datos de entorno y datos límites del medio, datos de uso y datos de inspección.

La figura 3.21 resume los procedimientos generales para insertar información al centro. La información requerida para la recuperación se extrae de los diferentes documentos de entrada y se almacena en cinta magnética, la cual conforma los archivos bibliográficos. Después de que la información de entrada ha sido codificada, los documentos son registrados en cinta video magnética, la cual conforma los archivos de

documentos. Entonces, los documentos originales son almacenados en un archivo manual, que consiste en gabinetes estándares de archivo.

El centro ofrece a sus usuarios dos servicios principales: las capacidades de un índice y la consulta. El índice es una publicación periódica que presenta de manera abstracta información para una parte o material específico. La función primordial del índice es proporcionarle al usuario un resumen de la información proveniente del centro, que se encuentra disponible. La capacidad de consulta es el principal servicio ofrecido por el centro, ya que le permite al usuario hacer consultas telefónicas al centro y recibir una respuesta ya sea telefónica o por correo.

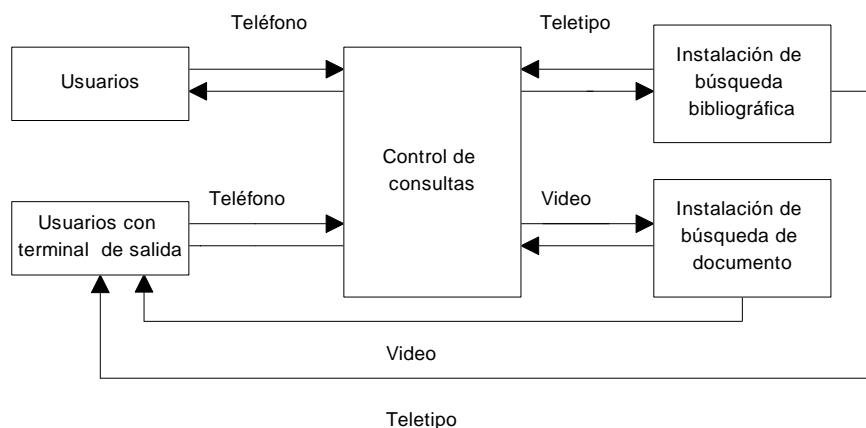


Figura 3.21: Diagrama de flujo generalizado de respuestas a consultas

Flujo de Procesamiento de Consultas:

Todas las consultas se reciben en la sección de control de consultas por vía telefónica, donde hay dos líneas telefónicas con dos teléfonos cada una. Debido a la complejidad de los tipos de consultas, no existe la posibilidad de que el usuario consulte el centro. En lugar de esto, el centro utiliza el personal técnico para que reciba, estandarice y codifique las consultas. Existen tres rutas de procesamientos según las clases de consultas: La búsqueda manual especial, la búsqueda bibliográfica y la búsqueda de documentos.

La búsqueda manual se lleva a cabo para las consultas que no pueden ser contestadas con los archivos bibliográficos.

La búsqueda bibliográfica es una búsqueda automática en los archivos bibliográficos para determinar, en forma resumida, la información que se encuentra disponible en el centro.

La búsqueda automática de documentos es una adición propuesta con respecto al centro. Esta clase de consulta es una petición de copias específicas de los documentos almacenados en los archivos de documentos.

Desarrollo del Modelo:

En la actualidad, el centro de información es operacional; sin embargo, fue esencial una evaluación para asegurarse de que el centro podía responder al incremento estimado del número de consultas. Además, numerosos estudios de equipo han resultado en las recomendaciones referentes a los tipos de hardware para la expansión del centro. Sin embargo, había un vacío en cuanto a la integración de estas recomendaciones en el centro actual, y luego en la evaluación de la operación del mismo.

El objetivo de este estudio era evaluar el centro actual, incluyendo las adiciones propuestas, para determinar la capacidad de las instalaciones del centro para recibir y responder a las consultas. El planteamiento seleccionado para evaluar el centro, fue determinar el tiempo de respuesta al servicio propuesto de consultas. El tiempo de respuesta se define como el tiempo total transcurrido, una vez que el solicitante ha entrado en contacto con el control de consultas a través del procesamiento de la consulta, hasta el momento en que la respuesta es remitida al solicitante.

Se usó una serie de instrucciones simples de cálculo para describir la respuesta de una consulta. La clasificación de estas instrucciones como componentes operativos y de demora, así como la suma de estos, dieron como resultado expresiones que representan el tiempo de respuesta. Una vez que se ha construido el modelo con las expresiones del tiempo de respuesta, el centro puede simularse bajo una variedad de condiciones. Debido a que estas condiciones eran controlables, el centro podía ser probado mediante la variación de un parámetro o combinación de parámetros. El modelo se basó en el siguiente conjunto de definiciones:

1. El procesamiento de una consulta requiere de la utilización de un conjunto definido de instalaciones (F_1, F_2, \dots, F_M).
2. Puede existir una cola de servicio (Q_1, Q_2, \dots, Q_N) entre las instalaciones.
3. Un componente (C_1, C_2, \dots, C_k) se define como una expresión de tiempo. El tiempo para procesar una consulta en la instalación se define como un componente operativo [$t_1(F_1), t_2(F_2), \dots, t_i(F_M)$]. El tiempo que puede transcurrir en una fila de espera se define como un componente de demora [$t_1(Q_1), t_2(Q_2), \dots, t_i(Q_N)$].
4. El tiempo de respuesta para cada tipo de consulta es una función de un conjunto de componentes (C_1, C_2, \dots, C_k) y puede expresarse como:

$$T_q = f(C_1, C_2, \dots, C_k) \quad (1)$$

El conjunto de componentes operativos y de demora asociados con cada T_q pueden expresarse como una función de:

$$T_q = f[t_1(F_1), t_2(F_2), \dots, t_i(F_M) \text{ y } t_1(Q_1), t_2(Q_2), \dots, t_i(Q_N)] \quad (2)$$

Asumiendo que los componentes operativos y de demora son independientes entre sí, podemos expresar T_q como:

$$T_q = \sum_{i=1}^M t_i(F_i) + \sum_{j=1}^N t_j(Q_j) \quad (3)$$

$$i = 1 \quad j = 1$$

donde T_q es el tiempo de respuesta para cada tipo de consulta ($q = 1, 2, \dots, 10$), $t_i(F_M)$ es el tiempo requerido en la instalación, M para procesar una consulta, y $t_j(Q_N)$ es el tiempo que puede transcurrir en la fila N antes de que se encuentre disponible la siguiente instalación para aceptar la consulta.

Recolección de Datos:

El esfuerzo de recolección de datos estaba orientado hacia la recolección de los datos representados por los componentes operativos, $t_i(F_M)$, y por las entradas adicionales requeridas para el programa de simulación. La recolección de datos se dividió en dos fases: La recolección de los datos empíricos disponibles, y la generación de los datos para los cuales no se encontraban disponibles los datos empíricos y que estaban asociados con las adiciones propuestas con respecto al centro.

Las fuentes principales de datos eran los registros de bitácora, grabaciones telefónicas e informes semanales y mensuales del estado interno. Se determinó a través de la prueba estadística que la distribución de las llegadas de las consultas entrantes era de tipo de Poisson. Debido a que las llegadas eran de tipo Poisson, el tiempo entre las consultas siguió la distribución exponencial negativa. Por consiguiente, las distribuciones exponenciales adecuadas se generaron a partir de las tasas de crecimiento proyectadas de las consultas. Se determinó que varias distribuciones de tiempo de servicio eran de tipo Erlang, tales como el tiempo que se empleaba en el teléfono y el tiempo en que se conduce una búsqueda bibliográfica. Varias de las distribuciones no podían ser aproximadas por cualquiera de las distribuciones teóricas estándar. Algunas de estas eran el número de preguntas por consultas de la clase 1, el número de documentos por respuestas y el número de páginas por documento. En estos ejemplos, los datos reales de los muestreros se usaron como entrada.

Planeación Táctica:

Al probar las capacidades de respuesta a consultas del centro, se puede considerar una combinación de parámetros como variables de entrada para la simulación, tales como los flujos proyectados de llegada de consultas. Por consiguiente, las instalaciones y las filas de esperas del centro se evalúan como una función de las varias tasas de flujos de llegada. Antes de ejecutar el modelo de simulación se tuvo que determinar las condiciones de arranque, la definición del equilibrio, las condiciones de paradas y el tamaño de la muestra.

Debido a que el modelo de simulación requiere que transcurra una cierta cantidad de tiempo de operación para que el sistema se acerque a su estado de equilibrio o estabilización, debe emitirse una parte inicial de la ejecución a partir de los resultados obtenidos en la salida. Este período de estabilización es una función del número de consultas que se encuentran en el centro y del número de consultas que éste ha procesado. De acuerdo con R.L. Conway, no hay criterios objetivos para determinar

cuándo debería empezar a recolectarse las mediciones de los resultados; sin embargo, recomienda el truncamiento de una serie de medidas hasta que una medida no sea el máximo ni el mínimo de las mediciones anteriores.

El tiempo requerido para que el sistema se acerque al equilibrio depende de la condición inicial de arranque. Debido a que el modelo comienza con todas sus instalaciones y filas inactivas y vacías, cualquier condición de arranque que no se encuentre inactiva y vacía disminuiría el tiempo requerido para que el sistema alcance el equilibrio. En forma ideal, la condición de arranque debería seleccionarse de manera que correspondiera con la condición cuando el centro haya alcanzado el equilibrio. Sin embargo, debido a que no se tiene dicho conocimiento en el centro, la selección de la condición de arranque debe ser subjetiva. La condición de arranque seleccionada por el centro inicialmente iba a generar 10 llegadas en el tiempo cero o inicial.

Después de que se generan estas llegadas, se usan las tasas seleccionadas de llegadas de entradas. El planteamiento adoptado por Conway se usa para determinar el punto de estabilización del centro. Las consultas se agrupan en conjuntos de 25 en el orden de su terminación con respecto al centro. El tiempo medio para cada conjunto se computa y se presenta en la figura 3.23. Los resultados que se muestran en ésta también muestran las condiciones de arranque previamente establecidas. La primera media, que no es ni el máximo, ni el mínimo de las series anteriores, se encuentra en el quinto conjunto. Por consiguiente, no fueron incluidos en los resultados de salida, los datos para las primeras 125 consultas. Este punto de estabilización se mantuvo constante en todas las ejecuciones.

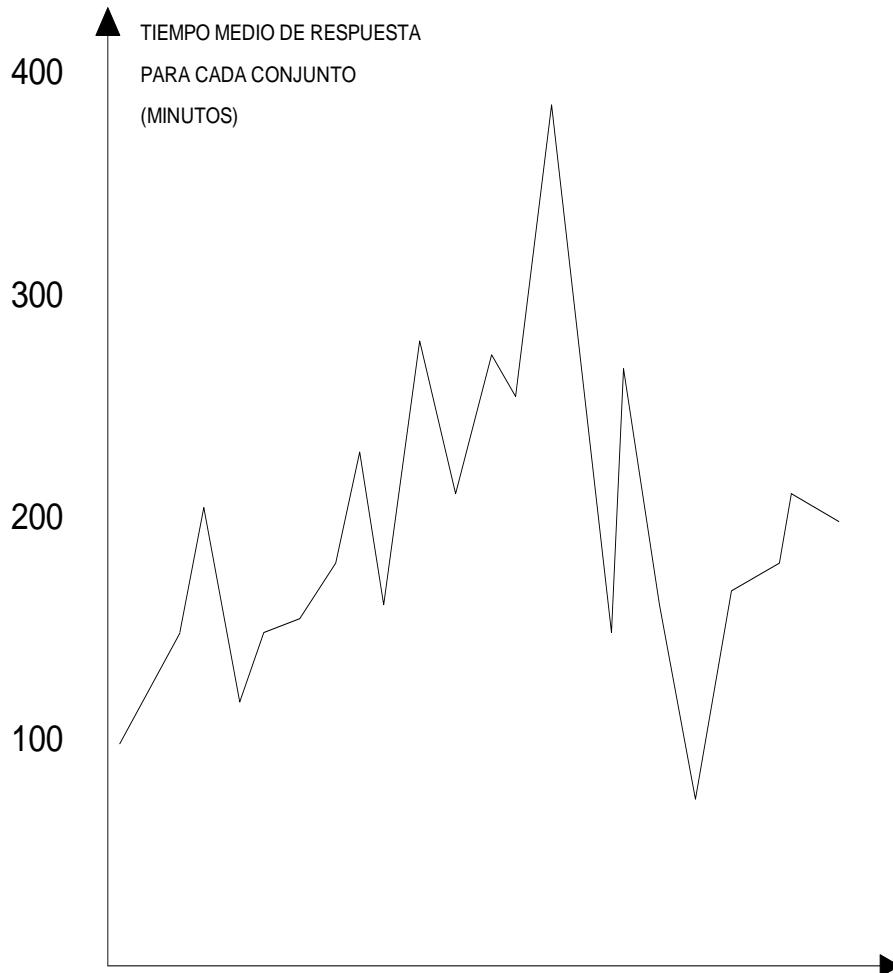


Figura 3.22: Tiempo medio de respuesta de consulta para cada conjunto de 25 consultas.

La condición de parada del modelo es la de detener la ejecución después de que un número determinado de consultas haya sido procesado por el centro. No se asigna ningún máximo al número de consultas insertadas en el modelo; por lo tanto, después que se obtiene la condición de parada, todavía pueden encontrarse algunas consultas que están siendo atendidas en las instalaciones o siendo demoradas en las filas de servicios.

Se usan muchos métodos para determinar los tamaños de la muestra. Una técnica que a menudo se usa es determinar el tamaño de la muestra que se necesita para obtener cierta confianza de que una instalación o fila se encuentre en un estado o condición determinada. Debido a que el centro tiene muchas instalaciones y filas, y debido a que éstas no se usan con una gran frecuencia, dicho planteamiento requeriría de una muestra muy grande. Consecuentemente, el tiempo de respuesta de la consulta se usó para determinar el tamaño de la muestra. Entonces, se usó una comparación relativa de las alternativas entre los resultados de las instalaciones y las filas.

A fin de determinar el tamaño de la muestra, se agruparon las consultas en conjuntos de 25 en el orden de su terminación con respecto al centro. Al computar el tiempo medio de respuesta para cada conjunto se obtuvo un grupo de n mediciones, $x_1, x_2, x_3, \dots, x_n$. Si se considera el teorema de límite central, es posible asumir que la x de cada grupo está distribuida normalmente. Mediante la desigualdad de Tchebycheff, nx puede ser escogida de tal modo que para un nivel de 95% de confianza, donde x es la estimación de la medida real, $(1.96) s_x = \sqrt{x}$ (4)

debido a que n mediciones están relacionadas con el tiempo, no es posible asumir que las mediciones consecutivas son independientes. Por consiguiente, se usó la varianza de la media de una serie auto correlacionada. Como resultados de los cálculos anteriores, se obtuvieron un tamaño mínimo de la muestra de 500.

Salidas del Modelo:

Las salidas de los modelos proporcionan información con respecto a los siguientes aspectos del centro:

- El tiempo medio para contestar los varios tipos de consultas (por ejemplo, el tiempo de respuesta).
- La utilización promedio de las instalaciones del centro.
- Los tiempos promedios de retardo en las instalaciones del centro.

Para complementar las salidas anteriores al modelo también son salidas:

- La distribución de los tiempos de respuesta.
- La distribución de los tiempos de demora en las filas.
- La longitud máxima de la fila.
- Las longitudes promedio de la fila.
- Las entradas totales que pasaron un tiempo cero en las colas.

El análisis de los resultados de las ejecuciones de simulación indica que, si se obtuvieran las tasas proyectadas de llegadas, el centro solamente tendría la necesidad de hacer algunas modificaciones menores referente a la configuración del equipo y a la asignación del personal. La mayoría de estas modificaciones sólo requieren de una pieza adicional de hardware o de un incremento o reubicación del personal asignado a las varias instalaciones. Este caso nos convence de que es necesario conocer muy bien los conceptos de distribución de probabilidades, pruebas de hipótesis, etc. Por tales motivos, los dos siguientes capítulos repasan los conceptos más importantes.

3.12 Estructura de los Modelos de Simulación

Antes de iniciar el proceso de desarrollo de un modelo, se debería comprender la estructura básica a partir de la cual se construyen los modelos. Aunque un modelo puede ser muy complicado matemática o físicamente, su estructura fundamental es muy simple.

En una perspectiva un poco más amplia, encontramos que casi cualquier modelo consiste de alguna combinación de los siguientes elementos:

- **Componentes:** Por componentes entendemos las partes constituyentes que en conjunto forman el sistema. Algunas veces, también nos referimos a los componentes como elementos o subsistemas.
- **Variables:** Sólo pueden suponer aquellos valores que la forma de la función permite. Las variables pueden ser cualquiera entidad o atributo de la entidad. Podemos reconocer dos tipos de variables en un modelo de un sistema, las exógenas y las endógenas. Las variables exógenas también se llaman variables de entrada, es decir, éstas son variables que se originan o se producen fuera del sistema o que surgen debido a causas externas. Las variables endógenas son aquellas producidas dentro del sistema o que resultan de causas internas. También podemos referirnos a las variables endógenas, ya sea como variables de estado o variables de salida. Los estadísticos llaman a las variables exógenas independientes y a las variables endógenas dependientes.
- **Parámetros:** Son cantidades a las cuales el operador del modelo puede asignarles valores arbitrarios, a diferencia de las variables. Otra manera de ver esto es que los parámetros, una vez establecidos, son *constantes* y no varían. Con frecuencia, el análisis estadístico implica intentos para determinar estos parámetros desconocidos, pero fijos, para un conjunto de datos.
- **Relaciones funcionales:** Describen a las variables y a los parámetros de tal manera que muestran su comportamiento dentro de un componente o entre componentes de un sistema.
- **Restricciones:** Son limitaciones impuestas a los valores de las variables o a la manera en la cual los recursos pueden asignarse o consumirse. Estas restricciones pueden ser auto-impuestas por el modelador o impuestas por el sistema mediante la naturaleza del mismo.
- **Objetivo:** Es una definición explícita de los objetivos o metas del sistema y de cómo se evaluarán. Por lo general, el objetivo es una parte integral del modelo y la manipulación del mismo se logra por los intentos para optimizar o satisfacer los criterios establecidos.

3.13 Resumen

Este capítulo presentó los conceptos más relevantes para poder entender la importancia de la modelación de sistemas. Como se puede concluir, hoy día se trabaja con modelos y no con sistemas. Ahora bien, para poder construir un modelo se requiere de una metodología oportuna, como la que se presentó, para poder construir modelos que se asemejen casi al 100% a las características de los sistemas. De otra forma, si el modelo no es una réplica, entonces la experimentación que se realice arrojará resultados muy cuestionables.

INTRODUCTION TO MODELING AND SIMULATION

Anu Maria

State University of New York at Binghamton
Department of Systems Science and Industrial Engineering
Binghamton, NY 13902-6000, U.S.A.

ABSTRACT

This introductory tutorial is an overview of simulation modeling and analysis. Many critical questions are answered in the paper. What is modeling? What is simulation? What is simulation modeling and analysis? What types of problems are suitable for simulation? How to select simulation software? What are the benefits and pitfalls in modeling and simulation? The intended audience is those unfamiliar with the area of discrete event simulation as well as beginners looking for an overview of the area. This includes anyone who is involved in system design and modification - system analysts, management personnel, engineers, military planners, economists, banking analysts, and computer scientists. Familiarity with probability and statistics is assumed.

1 WHAT IS MODELING?

Modeling is the process of producing a model; a model is a representation of the construction and working of some system of interest. A model is similar to but simpler than the system it represents. One purpose of a model is to enable the analyst to predict the effect of changes to the system. On the one hand, a model should be a close approximation to the real system and incorporate most of its salient features. On the other hand, it should not be so complex that it is impossible to understand and experiment with it. A good model is a judicious tradeoff between realism and simplicity. Simulation practitioners recommend increasing the complexity of a model iteratively. An important issue in modeling is model validity. Model validation techniques include simulating the model under known input conditions and comparing model output with system output.

Generally, a model intended for a simulation study is a mathematical model developed with the help of simulation software. Mathematical model classifications include deterministic (input and output variables are fixed values) or stochastic (at least one of the input or

output variables is probabilistic); static (time is not taken into account) or dynamic (time-varying interactions among variables are taken into account). Typically, simulation models are stochastic and dynamic.

2 WHAT IS SIMULATION?

A simulation of a system is the operation of a model of the system. The model can be reconfigured and experimented with; usually, this is impossible, too expensive or impractical to do in the system it represents. The operation of the model can be studied, and hence, properties concerning the behavior of the actual system or its subsystem can be inferred. In its broadest sense, simulation is a tool to evaluate the performance of a system, existing or proposed, under different configurations of interest and over long periods of real time.

Simulation is used before an existing system is altered or a new system built, to reduce the chances of failure to meet specifications, to eliminate unforeseen bottlenecks, to prevent under or over-utilization of resources, and to optimize system performance. For instance, simulation can be used to answer questions like: What is the best design for a new telecommunications network? What are the associated resource requirements? How will a telecommunication network perform when the traffic load increases by 50%? How will a new routing algorithm affect its performance? Which network protocol optimizes network performance? What will be the impact of a link failure?

The subject of this tutorial is *discrete event simulation* in which the central assumption is that the system changes instantaneously in response to certain discrete events. For instance, in an M/M/1 queue - a single server queuing process in which time between arrivals and service time are exponential - an arrival causes the system to change instantaneously. On the other hand, continuous simulators, like flight simulators and weather simulators, attempt to quantify the changes in a system continuously over time in response to

controls. Discrete event simulation is less detailed (coarser in its smallest time unit) than continuous simulation but it is much simpler to implement, and hence, is used in a wide variety of situations.

Figure 1 is a schematic of a simulation study. The iterative nature of the process is indicated by the system under study becoming the altered system which then becomes the system under study and the cycle repeats. In a simulation study, human decision making is required at all stages, namely, model development, experiment design, output analysis, conclusion formulation, and making decisions to alter the system under study. The only stage where human intervention is not required is the running of the simulations, which most simulation software packages perform efficiently. The important point is that powerful simulation software is merely a hygiene factor - its absence can hurt a simulation study but its presence will not ensure success. Experienced problem formulators and simulation modelers and analysts are indispensable for a successful simulation study.

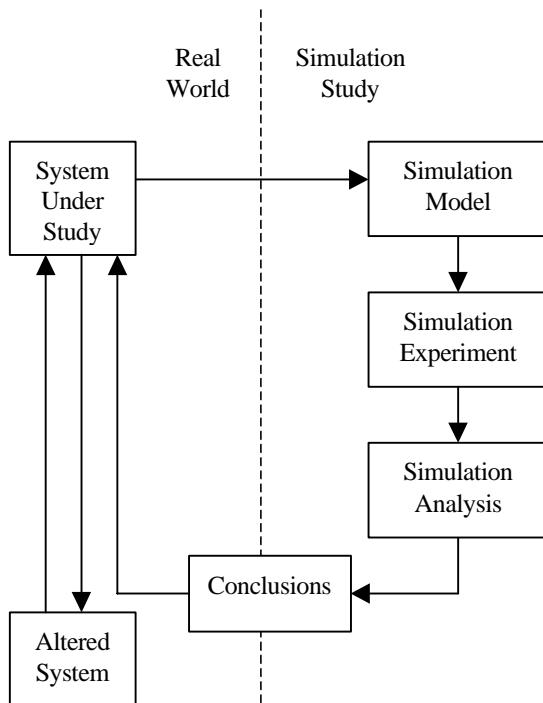


Figure 1: Simulation Study Schematic

The steps involved in developing a simulation model, designing a simulation experiment, and performing simulation analysis are:

- Step 1. Identify the problem.
- Step 2. Formulate the problem.
- Step 3. Collect and process real system data.
- Step 4. Formulate and develop a model.

- Step 5. Validate the model.
- Step 6. Document model for future use.
- Step 7. Select appropriate experimental design.
- Step 8. Establish experimental conditions for runs.
- Step 9. Perform simulation runs.
- Step 10. Interpret and present results.
- Step 11. Recommend further course of action.

Although this is a logical ordering of steps in a simulation study, many iterations at various sub-stages may be required before the objectives of a simulation study are achieved. Not all the steps may be possible and/or required. On the other hand, additional steps may have to be performed. The next three sections describe these steps in detail.

3 HOW TO DEVELOP A SIMULATION MODEL?

Simulation models consist of the following components: system entities, input variables, performance measures, and functional relationships. For instance in a simulation model of an M/M/1 queue, the server and the queue are system entities, arrival rate and service rate are input variables, mean wait time and maximum queue length are performance measures, and 'time in system = wait time + service time' is an example of a functional relationship. Almost all simulation software packages provide constructs to model each of the above components. Modeling is arguably the most important part of a simulation study. Indeed, a simulation study is as good as the simulation model. Simulation modeling comprises the following steps:

Step 1. Identify the problem. Enumerate problems with an existing system. Produce requirements for a proposed system.

Step 2. Formulate the problem. Select the bounds of the system, the problem or a part thereof, to be studied. Define overall objective of the study and a few specific issues to be addressed. Define performance measures - quantitative criteria on the basis of which different system configurations will be compared and ranked. Identify, briefly at this stage, the configurations of interest and formulate hypotheses about system performance. Decide the time frame of the study, i.e., will the model be used for a one-time decision (e.g., capital expenditure) or over a period of time on a regular basis (e.g., air traffic scheduling). Identify the end user of the simulation model, e.g., corporate management versus a production supervisor. Problems must be formulated as precisely as possible.

Step 3. Collect and process real system data. Collect data on system specifications (e.g., bandwidth for a communication network), input variables, as well as

performance of the existing system. Identify sources of randomness in the system, i.e., the stochastic input variables. Select an appropriate input probability distribution for each stochastic input variable and estimate corresponding parameter(s).

Software packages for distribution fitting and selection include ExpertFit, BestFit, and add-ons in some standard statistical packages. These aids combine goodness-of-fit tests, e.g., χ^2 test, Kolmogorov-Smirnov test, and Anderson-Darling test, and parameter estimation in a user friendly format.

Standard distributions, e.g., exponential, Poisson, normal, hyperexponential, etc., are easy to model and simulate. Although most simulation software packages include many distributions as a standard feature, issues relating to random number generators and generating random variates from various distributions are pertinent and should be looked into. Empirical distributions are used when standard distributions are not appropriate or do not fit the available system data. Triangular, uniform or normal distribution is used as a first guess when no data are available. For a detailed treatment of probability distributions see Maria and Zhang (1997).

Step 4. Formulate and develop a model. Develop schematics and network diagrams of the system (How do entities flow through the system?). Translate these conceptual models to simulation software acceptable form. Verify that the simulation model executes as intended. Verification techniques include traces, varying input parameters over their acceptable range and checking the output, substituting constants for random variables and manually checking results, and animation.

Step 5. Validate the model. Compare the model's performance under known conditions with the performance of the real system. Perform statistical inference tests and get the model examined by system experts. Assess the confidence that the end user places on the model and address problems if any. For major simulation studies, experienced consultants advocate a structured presentation of the model by the simulation analyst(s) before an audience of management and system experts. This not only ensures that the model assumptions are correct, complete and consistent, but also enhances confidence in the model.

Step 6. Document model for future use. Document objectives, assumptions and input variables in detail.

4 HOW TO DESIGN A SIMULATION EXPERIMENT?

A simulation experiment is a test or a series of tests in which meaningful changes are made to the input

variables of a simulation model so that we may observe and identify the reasons for changes in the performance measures. The number of experiments in a simulation study is greater than or equal to the number of questions being asked about the model (e.g., Is there a significant difference between the mean delay in communication networks A and B?, Which network has the least delay: A, B, or C? How will a new routing algorithm affect the performance of network B?). Design of a simulation experiment involves answering the question: what data need to be obtained, in what form, and how much? The following steps illustrate the process of designing a simulation experiment.

Step 7. Select appropriate experimental design. Select a performance measure, a few input variables that are likely to influence it, and the levels of each input variable. When the number of possible configurations (product of the number of input variables and the levels of each input variable) is large and the simulation model is complex, common second-order design classes including central composite, Box-Behnken, and full-factorial should be considered. Document the experimental design.

Step 8. Establish experimental conditions for runs. Address the question of obtaining accurate information and the most information from each run. Determine if the system is stationary (performance measure does not change over time) or non-stationary (performance measure changes over time). Generally, in stationary systems, steady-state behavior of the response variable is of interest. Ascertain whether a terminating or a non-terminating simulation run is appropriate. Select the run length. Select appropriate starting conditions (e.g., empty and idle, five customers in queue at time 0). Select the length of the warm-up period, if required. Decide the number of independent runs - each run uses a different random number stream and the same starting conditions - by considering output data sample size. Sample size must be large enough (at least 3-5 runs for each configuration) to provide the required confidence in the performance measure estimates. Alternately, use common random numbers to compare alternative configurations by using a separate random number stream for each sampling process in a configuration. Identify output data most likely to be correlated.

Step 9. Perform simulation runs. Perform runs according to steps 7-8 above.

5 HOW TO PERFORM SIMULATION ANALYSIS?

Most simulation packages provide run statistics (mean, standard deviation, minimum value, maximum value) on the performance measures, e.g., wait time (non-time persistent statistic), inventory on hand (time persistent statistic). Let the mean wait time in an M/M/1 queue observed from n runs be W_1, W_2, \dots, W_n . It is important to understand that the mean wait time W is a random variable and the objective of output analysis is to estimate the true mean of W and to quantify its variability.

Notwithstanding the facts that there are no data collection errors in simulation, the underlying model is fully known, and replications and configurations are user controlled, simulation results are difficult to interpret. An observation may be due to system characteristics or just a random occurrence. Normally, statistical inference can assess the significance of an observed phenomenon, but most statistical inference techniques assume independent, identically distributed (iid) data. Most types of simulation data are autocorrelated, and hence, do not satisfy this assumption. Analysis of simulation output data consists of the following steps.

Step 10. Interpret and present results. Compute numerical estimates (e.g., mean, confidence intervals) of the desired performance measure for each configuration of interest. To obtain confidence intervals for the mean of autocorrelated data, the technique of batch means can be used. In batch means, original contiguous data set from a run is replaced with a smaller data set containing the means of contiguous batches of original observations. The assumption that batch means are independent may not always be true; increasing total sample size and increasing the batch length may help.

Test hypotheses about system performance. Construct graphical displays (e.g., pie charts, histograms) of the output data. Document results and conclusions.

Step 11. Recommend further course of action. This may include further experiments to increase the precision and reduce the bias of estimators, to perform sensitivity analyses, etc.

6 AN EXAMPLE

A machine shop contains two drills, one straightener, and one finishing operator. Figure 2 shows a schematic of the machine shop. Two types of parts enter the machine shop.

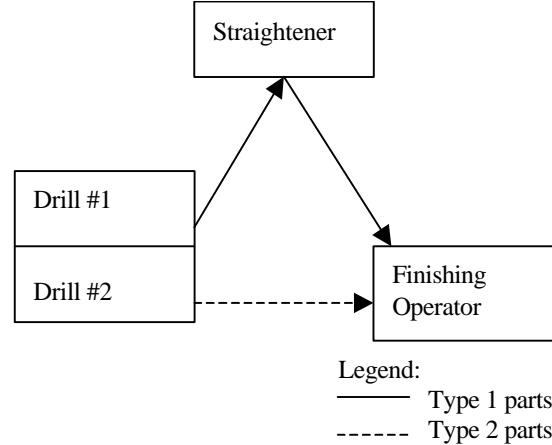


Figure 2: Schematic of the Machine Shop

Type 1 parts require drilling, straightening, and finishing in sequence. Type 2 parts require only drilling and finishing. The frequency of arrival and the time to be routed to the drilling area are deterministic for both types of parts.

Step 1. Identify the problem. The utilization of drills, straightener, and finishing operator needs to be assessed. In addition, the following modification to the original system is of interest: the frequency of arrival of both parts is exponential with the same respective means as in the original system.

Step 2. Formulate the problem. The objective is to obtain the utilization of drills, straightener, and finishing operator for the *original system* and the *modification*. The assumptions include:

- ◆ The two drills are identical
- ◆ There is no material handling time between the three operations.
- ◆ Machine availability implies operator availability.
- ◆ Parts are processed on a FIFO basis.
- ◆ All times are in minutes.

Step 3. Collect and process real system data. At the job shop, a Type 1 part arrives every 30 minutes, and a Type 2 part arrives every 20 minutes. It takes 2 minutes to route a Type 1 part and 10 minutes to route a Type 2 part to the drilling area. Parts wait in a queue till one of the two drilling machines becomes available. After drilling, Type 1 parts are routed to the straightener and Type 2 parts are

routed to the finishing operator. After straightening, Type 1 parts are routed to the finishing operator.

The operation times for either part were determined to be as follows. Drilling time is normally distributed with mean 10.0 and standard deviation 1.0. Straightening time is exponentially distributed with a mean of 15.0. Finishing requires 5 minutes per part.

Step 4. Formulate and develop a model. A model of the system and the modification was developed using a simulation package. A trace verified that the parts flowed through the job shop as expected.

Step 5. Validate the model. The utilization for a sufficiently long run of the original system was judged to be reasonable by the machine shop operators.

Step 6. Document model for future use. The models of the original system and the modification were documented as thoroughly as possible.

Step 7. Select appropriate experimental design. The original system and the modification described above were studied.

Step 8. Establish experimental conditions for runs. Each model was run three times for 4000 minutes and statistical registers were cleared at time 1000, so the statistics below were collected on the time interval [1000, 4000]. At the beginning of a simulation run, there were no parts in the machine shop.

Step 9. Perform simulation runs. Runs were performed as specified in Step 8 above.

Step 10. Interpret and present results. Table 1 contains the utilization statistics of the three operations for the original system and the modification (in parentheses).

Table 1: Utilization Statistics

	Drilling	Straightening	Finishing
Mean Run #1	0.83 (0.78)	0.51 (0.58)	0.42 (0.39)
Mean Run #2	0.82 (0.90)	0.52 (0.49)	0.41 (0.45)
Mean Run #3	0.84 (0.81)	0.42 (0.56)	0.42 (0.40)
Std. Dev. Run #1	0.69 (0.75)	0.50 (0.49)	0.49 (0.49)
Std. Dev. Run #2	0.68 (0.78)	0.50 (0.50)	0.49 (0.50)
Std. Dev. Run #3	0.69 (0.76)	0.49 (0.50)	0.49 (0.49)

Mean utilization represents the fraction of time a server is busy, i.e., busy time/total time. Furthermore, the average utilization output for drilling must be divided by the number of drills in order to get the utilization per drill. Each drill is busy about 40% of the time and straightening and finishing operations are busy about half the time. This implies that for the given work load, the system is

underutilized. Consequently, the average utilization did not change substantially between the original system and the modification; the standard deviation of the drilling operation seems to have increased because of the increased randomness in the modification. The statistical significance of these observations can be determined by computing confidence intervals on the mean utilization of the original and modified systems.

Step 11. Recommend further course of action. Other performance measures of interest may be: throughput of parts for the system, mean time in system for both types of parts, average and maximum queue lengths for each operation. Other modifications of interest may be: the flow of parts to the machine shop doubles, the finishing operation will be repeated for 10% of the products on a probabilistic basis.

7 WHAT MAKES A PROBLEM SUITABLE FOR SIMULATION MODELING AND ANALYSIS?

In general, whenever there is a need to model and analyze randomness in a system, simulation is the tool of choice. More specifically, situations in which simulation modeling and analysis is used include the following:

- ◆ It is impossible or extremely expensive to observe certain processes in the real world, e.g., next year's cancer statistics, performance of the next space shuttle, and the effect of Internet advertising on a company's sales.
- ◆ Problems in which mathematical model can be formulated but analytic solutions are either impossible (e.g., job shop scheduling problem, high-order difference equations) or too complicated (e.g., complex systems like the stock market, and large scale queuing models).
- ◆ It is impossible or extremely expensive to validate the mathematical model describing the system, e.g., due to insufficient data.

Applications of simulation abound in the areas of government, defense, computer and communication systems, manufacturing, transportation (air traffic control), health care, ecology and environment, sociological and behavioral studies, biosciences, epidemiology, services (bank teller scheduling), economics and business analysis.

8 HOW TO SELECT SIMULATION SOFTWARE?

Although a simulation model can be built using general purpose programming languages which are familiar to the analyst, available over a wide variety of platforms, and less expensive, most simulation studies today are implemented using a simulation package. The

advantages are reduced programming requirements; natural framework for simulation modeling; conceptual guidance; automated gathering of statistics; graphic symbolism for communication; animation; and increasingly, flexibility to change the model. There are hundreds of simulation products on the market, many with price tags of \$15,000 or more. Naturally, the question of how to select the best simulation software for an application arises. Metrics for evaluation include modeling flexibility, ease of use, modeling structure (hierarchical v/s flat; object-oriented v/s nested), code reusability, graphic user interface, animation, dynamic business graphics, hardware and software requirements, statistical capabilities, output reports and graphical plots, customer support, and documentation.

The two types of simulation packages are simulation languages and application-oriented simulators (Table 2). Simulation languages offer more flexibility than the application-oriented simulators. On the other hand, languages require varying amounts of programming expertise. Application-oriented simulators are easier to learn and have modeling constructs closely related to the application. Most simulation packages incorporate animation which is excellent for communication and can be used to debug the simulation program; a "correct looking" animation, however, is not a guarantee of a valid model. More importantly, animation is not a substitute for output analysis.

Table 2: Simulation Packages

Type Of Simulation Package	Examples
Simulation languages	Arena (previously SIMAN), AweSim! (previously SLAM II), Extend, GPSS, Micro Saint, SIMSCRIPT, SLX <i>Object-oriented software:</i> MODSIM III, SIMPLE++ <i>Animation software:</i> Proof Animation
Application -Oriented Simulators	<i>Manufacturing:</i> AutoMod, Extend+MFG, FACTOR/AIM, ManSim/X, MPSIM, ProModel, QUEST, Taylor II, WITNESS <i>Communications/computer:</i> COMNET III, NETWORK II.5, OPNET Modeler, OPNET Planner, SES/Strategizer, SES/workbench <i>Business:</i> BP\$IM, Extend+BPR, ProcessModel, ServiceModel, SIMPROCESS, Time machine <i>Health Care:</i> MedModel

9 BENEFITS OF SIMULATION MODELING AND ANALYSIS

According to practitioners, simulation modeling and analysis is one of the most frequently used operations research techniques. When used judiciously, simulation modeling and analysis makes it possible to:

- ◆ Obtain a better understanding of the system by developing a mathematical model of a system of

interest, and observing the system's operation in detail over long periods of time.

- ◆ Test hypotheses about the system for feasibility.
- ◆ Compress time to observe certain phenomena over long periods or expand time to observe a complex phenomenon in detail.
- ◆ Study the effects of certain informational, organizational, environmental and policy changes on the operation of a system by altering the system's model; this can be done without disrupting the real system and significantly reduces the risk of experimenting with the real system.
- ◆ Experiment with new or unknown situations about which only weak information is available.
- ◆ Identify the "driving" variables - ones that performance measures are most sensitive to - and the inter-relationships among them.
- ◆ Identify bottlenecks in the flow of entities (material, people, etc.) or information.
- ◆ Use multiple performance metrics for analyzing system configurations.
- ◆ Employ a systems approach to problem solving.
- ◆ Develop well designed and robust systems and reduce system development time.

10 WHAT ARE SOME PITFALLS TO GUARD AGAINST IN SIMULATION?

Simulation can be a time consuming and complex exercise, from modeling through output analysis, that necessitates the involvement of resident experts and decision makers in the entire process. Following is a checklist of pitfalls to guard against.

- ◆ Unclear objective.
- ◆ Using simulation when an analytic solution is appropriate.
- ◆ Invalid model.
- ◆ Simulation model too complex or too simple.
- ◆ Erroneous assumptions.
- ◆ Undocumented assumptions. This is extremely important and it is strongly suggested that assumptions made at each stage of the simulation modeling and analysis exercise be documented thoroughly.
- ◆ Using the wrong input probability distribution.
- ◆ Replacing a distribution (stochastic) by its mean (deterministic).
- ◆ Using the wrong performance measure.
- ◆ Bugs in the simulation program.
- ◆ Using standard statistical formulas that assume independence in simulation output analysis.
- ◆ Initial bias in output data.
- ◆ Making one simulation run for a configuration.

- ◆ Poor schedule and budget planning.
- ◆ Poor communication among the personnel involved in the simulation study.

REFERENCES

- Banks, J., J. S. Carson, II, and B. L. Nelson. 1996. *Discrete-Event System Simulation*, Second Edition, Prentice Hall.
- Bratley, P., B. L. Fox, and L. E. Schrage. 1987. *A Guide to Simulation*, Second Edition, Springer-Verlag.
- Fishwick, P. A. 1995. *Simulation Model Design and Execution: Building Digital Worlds*, Prentice-Hall.
- Freund, J. E. 1992. *Mathematical Statistics*, Fifth Edition, Prentice-Hall.
- Hogg, R. V., and A. T. Craig. 1995. *Introduction to Mathematical Statistics*, Fifth Edition, Prentice-Hall.
- Kleijnen, J. P. C. 1987. *Statistical Tools for Simulation Practitioners*, Marcel Dekker, New York.
- Law, A. M., and W. D. Kelton. 1991. *Simulation Modeling and Analysis*, Second Edition, McGraw-Hill.
- Law, A. M., and M. G. McComas. 1991. Secrets of Successful Simulation Studies, *Proceedings of the 1991 Winter Simulation Conference*, ed. J. M. Charnes, D. M. Morrice, D. T. Brunner, and J. J. Swain, 21-27. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.
- Maria, A., and L. Zhang. 1997. *Probability Distributions*, Version 1.0, July 1997, Monograph, Department of Systems Science and Industrial Engineering, SUNY at Binghamton, Binghamton, NY 13902.
- Montgomery, D. C. 1997. *Design and Analysis of Experiments*, Third Edition, John Wiley.
- Naylor, T. H., J. L. Balintfy, D. S. Burdick, and K. Chu. 1966. *Computer Simulation Techniques*, John Wiley.
- Nelson, B. L. 1995. *Stochastic Modeling: Analysis and Simulation*, McGraw-Hill.

AUTHOR BIOGRAPHY

ANU MARIA is an assistant professor in the department of Systems Science & Industrial Engineering at the State University of New York at Binghamton. She received her PhD in Industrial Engineering from the University of Oklahoma. Her research interests include optimizing the performance of materials used in electronic packaging (including solder paste, conductive adhesives, and underfills), simulation optimization techniques, genetics based algorithms for optimization of problems with a large number of continuous variables, multi criteria optimization, simulation, and interior-point methods.

1. La Simulación de Sistemas

1.1 Sistemas y Modelos

El término sistema se utiliza habitualmente con múltiples sentidos, tantos que resulta difícil dar una definición única que los abarque todos y al mismo tiempo sea lo suficientemente precisa para servir a propósitos específicos. Podemos partir de la definición de sistema como *conjunto de cosas que ordenadamente relacionadas entre si contribuyen a determinado objeto*. Se trata de una definición sencilla pero que pone de manifiesto los caracteres relevantes de lo que constituye el denominado enfoque sistémico: contemplación del todo y no de las partes aisladamente, acento en las relaciones entre las partes y consideración teleológica al tener en cuenta los propósitos u objetivos del sistema, especialmente válida para los sistemas creados por el hombre.

La orientación sistémica a la resolución de problemas reconoce que el comportamiento de cualquier parte tiene algún efecto sobre el comportamiento del sistema como un todo, en su desarrollo a partir de los años veinte ha ido solapando e interaccionando múltiples disciplinas, dando lugar a lo que hoy en día se conoce como **Investigación Operativa**, definida por Ackoff y Sasienski como “*la aplicación del método científico mediante equipos interprofesionales a los problemas de gobierno de sistemas organizados (hombre-máquina) para proporcionar soluciones que sirvan lo mejor posible a la organización considerada como un todo*”. Caracterizada metodológicamente por:

1. La aplicación del método científico a los problemas que se presentan en el diseño, construcción y gobierno de sistemas complejos en los que intervienen hombres y máquinas.
2. Enfoque global, es decir orientación sistémica
3. Construcción de modelos de los sistemas (Representación de los sistemas por medio de modelos)
4. Optimización: búsqueda de las mejores soluciones
5. Ayuda a los responsables de la gestión del sistema a la toma de decisiones

En el estudio de sistemas existentes, o en el diseño de sistemas que se han de construir, el primer paso lo constituye un proceso de adquisición de información sobre el sistema, qué elementos lo componen, o lo han de componer, qué propiedades los caracterizan, o los han de caracterizar, qué relaciones o interacciones existen, o se propone que existan, entre ellos, qué cambios provocan en el sistema tales interacciones, etc..

Denominaremos *entidades* a los objetos de interés que constituyen el sistema, *atributos* a las propiedades que caracterizan a las entidades componentes del sistema, y *estado del sistema* a la caracterización de las entidades del sistema y sus atributos en un instante dado. Nos va a interesar el carácter dinámico de los sistemas, es decir sus cambios de estado a lo largo del tiempo dentro de un horizonte dado, y en consecuencia nos va a interesar identificar qué es lo que produce cambios en el estado del sistema, y por lo tanto, estudiar la evolución del sistema a partir del seguimiento de sus cambios de estado.

La forma primaria de realizar este estudio sería, evidentemente, la experimentación con el propio sistema. Esto no siempre es posible. En unos casos por imposibilidad física o económica manifiesta, basta pensar en lo que implicaría el experimentar con una fábrica. En otros porque el sistema existe

únicamente en forma hipotética y precisamente lo que nos interesa es saber como se comportará antes de que sea construido.

Como consecuencia lo que haremos será estudiar el comportamiento del sistema a través de una representación formal, o modelo, del mismo.

1.2 El Proceso de Construcción de Modelos: Modelos Matemáticos

El análisis del sistema a través de un modelo implica que la representación del sistema que constituye el modelo ha de ser una representación manipulable. El ejercicio de construcción del modelo del sistema comienza por la construcción de un **modelo conceptual** del sistema, representación equivalente lógica aproximada del sistema real que como tal constituye una abstracción simplificada del mismo, que a continuación se traduce en un modelo apto para su ejecución en un ordenador. El proceso de modelización o construcción del modelo implica:

- Identificación de las entidades principales del sistema y de sus atributos característicos.
- Identificación y representación de las reglas que gobiernan el sistema que se quiere simular.
- Captación de la naturaleza de las interacciones lógicas del sistema que se modeliza.
- Verificación de que las reglas incorporadas al modelo son una representación válida de las del sistema que se modeliza.
- Formulación de hipótesis de modelización que traduzcan adecuadamente la información sobre el sistema y las relaciones entre sus entidades, y su papel en los cambios de estado del sistema
- Representación del comportamiento aleatorio inherente a las componentes del sistema que lo exhiban.

Las etapas del proceso de construcción del modelo se sintetizan en la figura 1, [GAR].

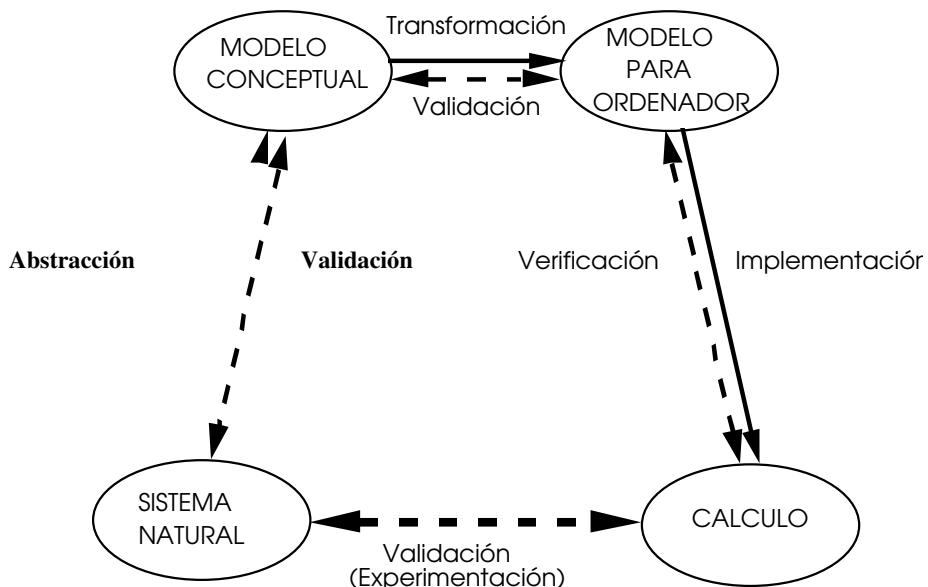


Figura 2. PROCESO DE MODELIZACION

Una precaución importante a tener en cuenta cuando se construye un modelo es que ningún modelo es mejor que las hipótesis que encierra. Traducir el modelo a un modelo específico para ordenador consiste en representar el modelo conceptual mediante un lenguaje apto para su ejecución en un ordenador. Este proceso se simplifica cuando la representación se hace utilizando un lenguaje especializado orientado a problemas específicos.

Tradicionalmente se han utilizado muchos tipos de modelos en el análisis de sistemas, clasificados de diferentes modos. Para los propósitos que nos interesan en esta descripción vamos a considerar únicamente los *modelos matemáticos* de los sistemas, es decir modelos en los que la representación formal del sistema que entendemos por modelo se hace en términos del formalismo de las matemáticas, los modelos matemáticos pueden ser a su vez *estáticos* o *dinámicos*. En el caso de los modelos matemáticos hemos de hacer referencia a la técnica utilizada para resolver el modelo, según la cual distinguiremos entre métodos *analíticos* y *numéricos*. **La simulación de sistemas es una técnica numérica que utiliza modelos matemáticos dinámicos.**

En un modelo matemático las entidades de un sistema y sus atributos se representan mediante variables matemáticas, clasificadas en *variables de control* y *variables no controlables*, según que los valores que puedan tomar puedan ser el resultado de una decisión o vengan determinados por las características del propio sistema. Las actividades que cambian el estado del sistema se describen por medio de funciones matemáticas que interrelacionan las variables. Los objetivos del sistema se representan mediante una *función de utilidad*, o *función objetivo*, que evalúa el rendimiento del sistema como una función de los dos tipos de variables controlables y no controlables.

La forma general de un modelo matemático propuesta por Ackoff, [ACK] es:

$$U = f(X, Y)$$

sometida a:

$$X \in W(Y)$$

donde U es la función de utilidad, dependiente de las variables de control $X=(X_1, X_2, \dots, X_n)$, y de la no controlables $Y=(Y_1, Y_2, \dots, Y_m)$, donde las variables de control pueden tomar valores en el dominio de definición $W(Y)$, que depende de las variables no controlables.

Puesto que en la mayor parte de los casos el objetivo que se persigue es el de determinar cuáles son las mejores decisiones, o en otras palabras, que valores de las variables de decisión optimizan la función de utilidad, y dado que como hemos dicho en los modelos matemáticos las relaciones entre las variables son funciones matemáticas, la forma genérica que adoptaremos para los modelos matemáticos es:

$$[OPT] U = f(X, Y)$$

sometida a:

$$R_k(X, Y) (\leq, =, \geq) b_k$$

$$k = 1, 2, \dots, K$$

donde $R_k(X, Y)$ es la k -ésima ecuación o inecuación de condición que traduce las relaciones funcionales entre las variables, o en otras palabras la formalización en términos de una ecuación de las hipótesis de modelización.

Un modelo matemático de tipo dinámico permite que los cambios en los atributos del sistema sean expresados como una función del tiempo, bien mediante una solución analítica o por medio de una computación numérica, según sea la complejidad del modelo. Supongamos, por ejemplo, el sistema correspondiente a la suspensión de una rueda de automóvil cuando se supone que la carrocería permanece inmóvil en la dirección vertical. El sistema puede ser representado como el caso de una masa M , sometida a una fuerza $F(t)$, que varía con el tiempo, y ligada a un muelle cuya fuerza es proporcional a su extensión o contracción, y al efecto de un absorbente de los impactos que ejerce una fuerza de amortiguación proporcional a la velocidad de la masa. El movimiento del sistema puede describirse por medio de la siguiente ecuación diferencial:

$$M\ddot{x} + D\dot{x} + Kx = KF(t)$$

donde x es la distancia que se ha desplazado, M es la masa, K es la constante elástica del muelle, y D es el factor de amortiguación de los impactos. Las componentes del sistema son en este caso la rueda, el muelle y el sistema amortiguador, y las hipótesis de modelización son las de la dinámica de un cuerpo sometido a la acción de fuerzas elásticas. Esta ecuación es un ejemplo de modelo matemático dinámico; una ecuación que en este caso puede ser resuelta analíticamente.

En otros casos la naturaleza de las ecuaciones que constituyen el modelo continuo hace aconsejable recurrir a procedimientos numéricos para su integración. Un ejemplo interesante de esta situación lo constituyen los modelos utilizados para estudiar la dinámica de poblaciones. Un caso típico es el de la dinámica de la competición entre dos poblaciones una pasiva, que sirve de presa, y otra activa,

depredadora de la anterior. Modelos de este tipo fueron introducidos por Volterra, [VOL], en 1926 como modelo biológico para estudiar la evolución de la población piscícola en el Adriático.

Los resultados de la integración del modelo dinámico presa-depredador de Lotka-Volterra, han sido obtenidos mediante el software numérico MATLAB, [MAT].

Otro ejemplo típico de modelos dinámicos son los Modelos Macroscópicos de Simulación de Tráfico. La perspectiva de la modelización macroscópica de los flujos de tráfico es la de adoptar un símil hidráulico, es decir visualizar el flujo de vehículos como flujo o corriente de un fluido continuo, es una asociación natural bastante intuitiva, que de hecho ya se adopta cuando se tiende a describir el tráfico en términos de volúmenes (o flujos), concentraciones (o densidades) y velocidades. En la analogía como flujo de un fluido, el tráfico es tratado como un fluido unidimensional compresible, lo que conduce a dos hipótesis básicas de modelización:

- a.** El flujo de tráfico se conserva
- b.** Hay una relación en términos de una función univaluada entre la velocidad y la densidad, o entre el flujo y la densidad.

En el modelo matemático la primera hipótesis se traduce en la denominada ecuación de conservación o de continuidad, que implica que si el flujo disminuye con la distancia la densidad se incrementa con el tiempo. En términos prácticos de ingeniería de tráfico la ecuación de conservación implica que en cualquier sistema de tráfico el flujo de entrada es igual al de salida más el almacenado en el sistema. Esta hipótesis se acepta, en general, sin mayor controversia, como un supuesto obvio.

No ocurre lo mismo, sin embargo, con la segunda hipótesis, que ha levantado y sigue levantando una gran controversia, en parte porque no siempre se entiende y porque las medidas son contradictorias. Concretamente si la velocidad (u) es una función de la densidad como consecuencia los conductores ajustan su velocidad de acuerdo con la densidad, es decir cuando la densidad aumenta la velocidad disminuye. Esto es intuitivamente correcto pero teóricamente puede conducir a velocidades o densidades negativas. Por otra parte, se ha observado que para el mismo valor de la densidad pueden medirse muchos valores de la velocidad. Todo ello conduce a la necesidad de matizar esta hipótesis. La matización consiste en considerar que la velocidad (o el flujo) es una función de la densidad únicamente en el equilibrio, y puesto que es muy raro poder observar el equilibrio es difícil obtener una relación velocidad-densidad satisfactoria, lo que lleva en la práctica a proponer una relación teórica.

1.3 Simulación de Sistemas Continuos y Simulación de Sistemas Discretos

En general los modelos matemáticos de tipo dinámico representan sistemas continuos, es decir sistemas en los que las actividades predominantes del sistema causan pequeños cambios en los atributos de sus entidades, cuando las relaciones entre ellas describen las tasas o ratios de cambio de los atributos, por lo que, en general, tales modelos están definidos formalmente por ecuaciones diferenciales.

En muchos casos a partir del modelo matemático del sistema es posible obtener información sobre el mismo por medios analíticos, como pone de manifiesto el modelo del sistema de amortiguación de un automóvil, que acabamos de presentar como ejemplo de modelo dinámico. Cuando esto no es posible se recurre a procedimientos numéricos para resolver las ecuaciones del modelo, especialmente en el caso de los modelos dinámicos representados por ecuaciones o sistemas de ecuaciones diferenciales.

Con el tiempo se ha ido desarrollando una amplia variedad de métodos numéricos de cálculo para resolver las ecuaciones de los modelos, una técnica numérica particular es la que denominamos *Simulación de Sistemas*, que como veremos consiste en un **seguimiento a lo largo del tiempo de los cambios que tienen lugar en el modelo dinámico del sistema**. En el caso de los modelos estáticos las técnicas numéricas involucradas se conocen como **Simulación Determinista**, mientras que en los modelos dinámicos bien por la naturaleza del sistema modelizado, bien por las características del proceso numérico utilizado, la introducción de la aleatoriedad nos llevará a hablar de **Simulación Estocástica**, que es la que va a ser objeto de este artículo.

La manera de efectuar el seguimiento temporal de los cambios en el modelo, que supondremos en correspondencia con los cambios en el sistema representado por el modelo, nos lleva a la aparición de dos grandes categorías dentro de la Simulación de Sistemas según que los cambios sean continuos o discretos. En el primer caso se supone que la naturaleza del sistema permite cambios de estado continuos, determinados por cambios continuos en los valores de las variables que representan el estado del sistema, mientras que en el segundo los cambios solo pueden tener lugar en instantes discretos en el tiempo.

Para los sistemas con cambios continuos, dado que nuestro principal interés a la hora de simular su comportamiento será reproducirlos, los sistemas de ecuaciones diferenciales serán la forma más adecuada de representarlos. Denominaremos *Simulación Continua* a la simulación basada en este tipo de modelos. Los simuladores analógicos han sido ampliamente utilizados en este tipo de simulación, aunque el desarrollo de las técnicas numéricas para la resolución de sistemas de ecuaciones diferenciales, el avance tecnológico en los ordenadores digitales, y la evolución de los lenguajes de programación les han hecho perder protagonismo. SIMULINK, software para la simulación de sistemas dinámicos integrado en el entorno de computación numérica MATLAB [MAT] es un buen ejemplo de esta tendencia.

Para los sistemas discretos, el seguimiento de los cambios de estado requiere la identificación de qué es lo que causa el cambio y cuando lo causa, lo que denominaremos un *suceso*, las ecuaciones del modelo se convierten entonces en las ecuaciones y relaciones lógicas que determinan las condiciones en que tiene lugar la ocurrencia de un suceso. Este tipo de simulación, conocida con el nombre de *Simulación Discreta*, consiste en el *seguimiento de los cambios de estado del sistema que tienen lugar como consecuencia de la ocurrencia de una secuencia de sucesos*.

1.4 La Simulación como Proceso Experimental: Experimentos y Ordenadores

La práctica de la simulación es una técnica que no realiza ningún intento específico para aislar las relaciones entre variables particulares, antes bien adopta un punto de vista global desde el que se intenta observar como cambian conjuntamente todas las variables del modelo con el tiempo. En todo caso, las relaciones entre las variables deben obtenerse a partir de tales observaciones. Esta concepción caracteriza la simulación como una técnica experimental de resolución de problemas, lo que comporta la necesidad de repetir múltiples ejecuciones de la simulación para poder entender las relaciones implicadas por el sistema, en consecuencia el uso de la simulación en un estudio debe planificarse como una serie de experimentos cuyo diseño debe seguir las normas del diseño de experimentos para que los resultados obtenidos puedan conducir a interpretaciones significativas de las relaciones de interés.

La simulación con computador es por lo tanto una técnica que realiza experimentos en un computador con un modelo de un sistema dado. El modelo es el vehículo utilizado para la experimentación en sustitución del sistema real. Los experimentos pueden llegar a tener un alto grado de sofisticación que requiera la utilización de técnicas estadísticas de diseño de experimentos. En la mayor parte de los casos los experimentos de simulación son la manera de obtener respuestas a preguntas del tipo ¿Qué pasaría si?, preguntas cuyo objetivo suele ser evaluar el impacto de una posible alternativa que sirva de soporte a un proceso de toma de decisiones sobre un sistema, proceso que puede representarse esquemáticamente mediante el diagrama siguiente, figura 4, Pidd [PIDa]:



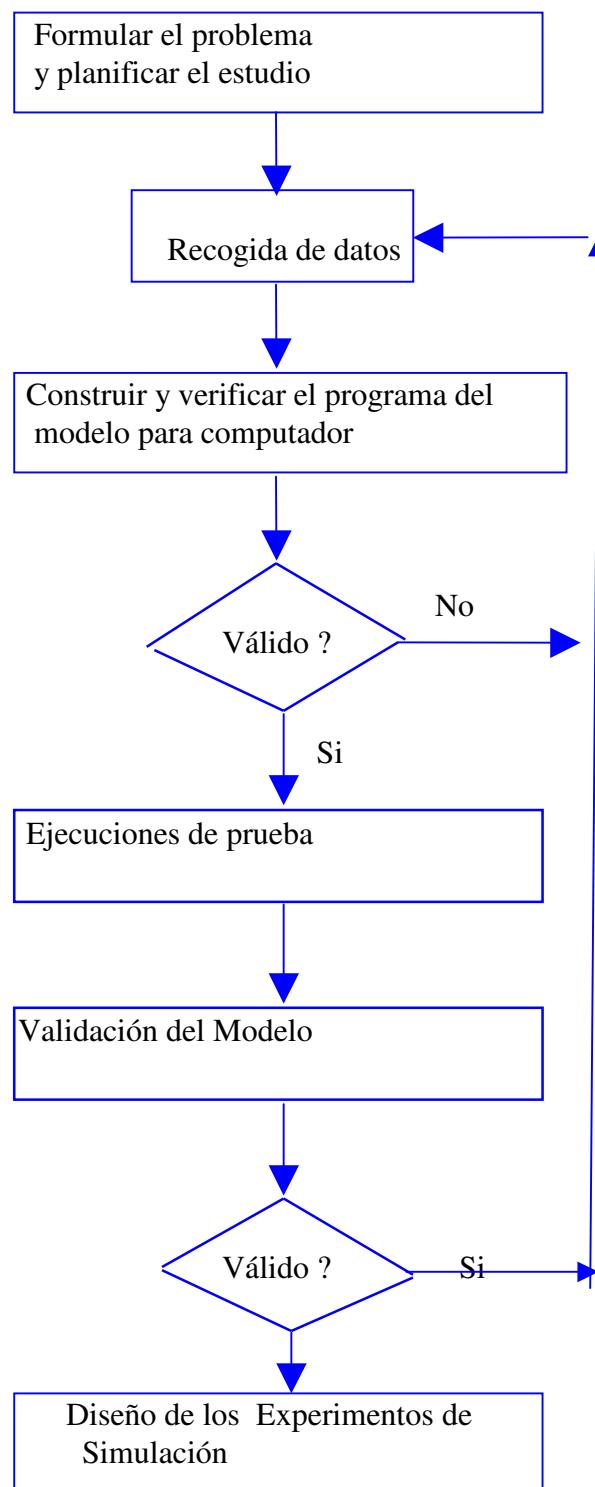
Figura 4. Esquema del proceso experimental de la simulación

La simulación, y los experimentos de simulación, se convierten así en una herramienta de **análisis de sistemas**, para entender cómo opera un sistema existente, o cómo puede operar uno propuesto. La situación ideal, en la cual el investigador realizaría los experimentos sobre el sistema real es sustituida por una en la que el investigador construye un modelo del sistema y experimenta sobre él mediante la simulación, utilizando un ordenador, para investigar el comportamiento del modelo e interpretar los resultados en términos del comportamiento del sistema objeto del estudio. La simulación, y el procedimiento experimental asociado, se convierten también en una herramienta de **diseño de sistemas**, cuyo objetivo es la producción de un sistema que satisfaga ciertas especificaciones. El diseñador puede seleccionar o planear como deben ser las componentes del sistema y concebir cual debe ser la combinación de componentes y relaciones entre ellas que determinan el sistema propuesto. El diseño se traduce en un modelo cuyo comportamiento permite inducir el del sistema previsto. El diseño se acepta cuando las previsiones se ajustan adecuadamente a los comportamientos deseados, en caso contrario se introducen las modificaciones pertinentes en el modelo y se repite el proceso.

Otra posibilidad es la que se da en estudios económicos, políticos, médicos, etc. en los que se conoce el comportamiento del sistema pero no los procesos que producen tal comportamiento. En este caso se formulan hipótesis sobre las entidades y actividades que pueden explicar la conducta. El estudio de simulación por medio del modelo correspondiente permite comparar las respuestas de un modelo basado en tales hipótesis con el comportamiento conocido, de manera que una concordancia adecuada lleva a suponer que la estructura del modelo se corresponde con la del sistema real. La aplicación de la simulación a diferentes tipos de sistemas combinada con las diferentes clases de estudio que se pueden realizar conduce a una gran cantidad de variantes de la manera en que se puede realizar un estudio de simulación. Sin embargo hay determinados pasos básicos del proceso que pueden identificarse como los constituyentes de lo que denominaremos la metodología de un estudio de simulación, y son los siguientes:

1. Definición del problema y planificación del estudio

2. Recogida de Datos
3. Formulación del modelo matemático
4. Construcción y verificación del programa para computador del modelo
5. Ejecuciones de prueba del modelo
6. Validación del modelo
7. Diseño de los experimentos de simulación
8. Ejecución de los experimentos.
9. Análisis de los resultados



Ejecución de los Experimentos de Simulación



Figura 5. Etapas de un estudio de Simulación

El proceso no es, en general, secuencial, sino iterativo, en el que algunos de los pasos pueden tener que repetirse en función de los resultados intermedios tal como muestra la figura 5.

Ningún estudio de simulación puede llevarse a cabo sin establecer claramente una definición precisa del problema que se pretende resolver y los objetivos del estudio. Los diseños alternativos del sistema que se han de estudiar han de quedar claramente especificados, así como los criterios para evaluar dichos diseños. Criterios que servirán de base al proceso de toma de decisiones para elegir uno de los diseños. Para la formulación del modelo debe establecerse su estructura definiendo cuales son los aspectos del funcionamiento del sistema que son significativos para la resolución del problema que tenemos entre manos, y que datos es necesario recoger para proporcionar al modelo la información adecuada.

La construcción del modelo de simulación es en muchos casos más un arte que una ciencia, que combina aspectos matemáticos y lógicos. En general la experiencia recomienda empezar con modelos moderadamente detallados que paulatinamente se van haciendo más sofisticados. El modelo únicamente debe contener el nivel de detalle requerido por los objetivos del estudio. Dado un modelo matemático la construcción del programa para computador es el requisito imprescindible para poder manipular numéricamente el modelo para obtener las soluciones que respondan a las preguntas que el analista se formula sobre el sistema.

La validación del modelo es uno de los pasos cruciales del proceso, suele ser uno de los más difíciles, pero es un requisito indispensable para establecer si el modelo representa o no adecuadamente el sistema objeto del estudio, de manera que se puedan garantizar las inducciones y extrapolaciones sobre el comportamiento del sistema a partir de lo observado sobre el modelo.

Diseñar los experimentos comporta, como hemos comentado anteriormente, aplicar rigurosamente las técnicas observacionales de la estadística, propias del método científico, que permitan garantizar la significación de las respuestas producidas por la ejecución del programa que implanta el modelo en el computador.

1.5 Modelos de Simulación frente a Soluciones Analíticas

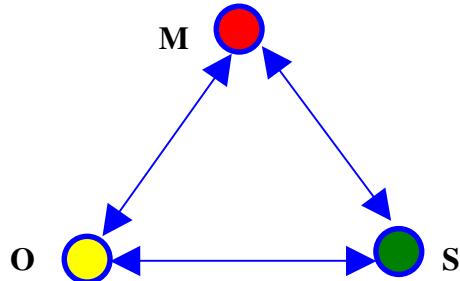
A pesar de su utilidad la simulación no puede considerarse como una panacea capaz de resolver todo tipo de situaciones, aún contando con la ayuda de los lenguajes especializados para la simulación, o de los avances que han representado los entornos software para simulación, Henriksson [HEN], el desarrollo de los generadores de simuladores, Mathewson [MTH], o de los simuladores visuales, SIMFACTORY [SMF], WITNESS [WIT], etc., la realización de un estudio de simulación puede comportar un esfuerzo y un consumo de recursos no despreciable en cualquiera de las fases: definición del problema, recogida de información, construcción del modelo y programación del mismo, realización de los experimentos de simulación en computador. Especialmente en este último caso sistemas complejos pueden conducir a

programas largos y complejos que requieran cantidades importantes de recursos computacionales. Estas han sido algunas de las razones por las que en ciertos dominios de aplicación la simulación ha sido considerada como un último recurso al que acudir cuando todo lo demás falla.

Si embargo la simulación, por sus características, y por los desarrollos computacionales que se han conseguido en los últimos años, sigue presentando una serie de ventajas que no solo la convierten en el procedimiento mas adecuado en muchos casos, sino que hacen que sea la única alternativa tecnológica en muchos otros.

Esto resulta especialmente obvio en aquellos casos en los que las características del sistema que se pretende estudiar hacen inviable, por razones físicas o de coste, la experimentación directa sobre el sistema. El mundo de la producción industrial, del tráfico, la aeronáutica, la industria del automóvil, etc. son claros ejemplos de esta situación, en la que, si bien es cierto que en algunos casos se puede recurrir a modelos analíticos, también lo es que tales modelos no siempre son capaces de recoger todos los aspectos de interés del sistema, que conducirían a modelos inviables, o para los que no se dispone de herramientas adecuadas, obligando a introducir una serie de hipótesis simplificadoras que pueden resultar inadecuadas en función de los objetivos del estudio. Un ejemplo relativamente sencillo puede ilustrar esta afirmación. Los modelos de colas son analíticamente tratables bajo hipótesis de modelización relativamente simples: llegadas según distribuciones de Poisson, tiempos de servicio exponenciales, disciplinas FIFO, etc.. Aún en este caso las soluciones para los períodos transitorios pueden ser complicadas de obtener analíticamente, en contraste con la simplicidad de los procedimientos para obtener las soluciones estacionarias, sin embargo, basta introducir hipótesis adicionales aparentemente sencillas, que aproximan el modelo a otras situaciones reales, para entrar rápidamente en el terreno de las dificultades analíticas crecientes, los modelos con distribuciones de probabilidad de llegadas y servicios de tipo general, inclusión de impaciencias o políticas basadas en prioridades, etc., son un buen ejemplo de ello. Incluso en aquellos casos en los que es posible la experimentación directa la simulación puede ofrecer ventajas tales como un coste inferior, tiempo, repeticiones y seguridad. Aún siendo viables los experimentos directos con el sistema físico pueden con frecuencia tener un coste muy superior al de la simulación a pesar de los esfuerzos para construir el modelo y el tiempo y recursos computacionales requeridos para la ejecución de los experimentos. Aunque el desarrollo de un modelo adecuado y su programación para ser ejecutado en un ordenador puede requerir una cantidad de tiempo significativa, una vez construido y depurado el modelo de simulación representa una atractiva posibilidad para trabajar con las más variadas escalas de tiempo, minutos, horas, semanas, meses, años, etc., en unos pocos segundos de tiempo de computador, lo que permite comparar colecciones variadas de alternativas, a través de experimentos de simulación que siempre pueden repetirse en las más diversas condiciones, lo que no siempre es posible en los experimentos con el sistema real, basta pensar en sistemas de manufactura, gestión de empresas, o militares (es difícil pensar en una situación en la que el enemigo permite repetir una batalla). Finalmente, es frecuente que los experimentos persigan el objetivo de determinar la respuesta del sistema en condiciones extremas, lo que puede resultar peligroso o incluso ilegal en la vida real. Las aplicaciones de la simulación en aeronáutica, o en la gestión de aeropuertos constituyen buenos ejemplos de lo que queremos significar. El dilema modelos analíticos frente a modelos de simulación debe resolverse en cada caso ateniéndose al tiempo de sistema, los objetivos del estudio, las características del modelo, los costes, etc.. La cuestión clave es nuestra habilidad y capacidad para construir el modelo del sistema, si este es analítico y las hipótesis de modelización no obligan a simplificaciones que invaliden la capacidad del modelo para responder a las cuestiones de interés que nos planteamos sobre el sistema, entonces las soluciones analíticas del modelo matemático pueden ser suficientes. Si nuestro conocimiento del sistema no nos permite formular hipótesis que conduzcan a una completa formalización del modelo en términos

analíticos, o el requisito de no realizar hipótesis simplificadoras conduce a modelos matemáticos de difícil, o imposible, tratamiento matemático, entonces posiblemente la simulación será la alternativa válida, si no es la única posible. El analista del sistema no debe olvidar que un mismo sistema puede representarse formalmente mediante diversos modelos en función de los problemas que el analista se plantea sobre el sistema. De acuerdo con Minsky, un objeto matemático M es un modelo válido de un sistema S para un observador O , si M es capaz de proporcionar respuestas válidas a las preguntas que O formula sobre S .



Law [LAW], formaliza el proceso de decisión modelo matemático-modelo de simulación, experimentación sobre el sistema real o sobre un modelo del sistema, como formas de estudiar un sistema, por medio del diagrama de la figura 6.

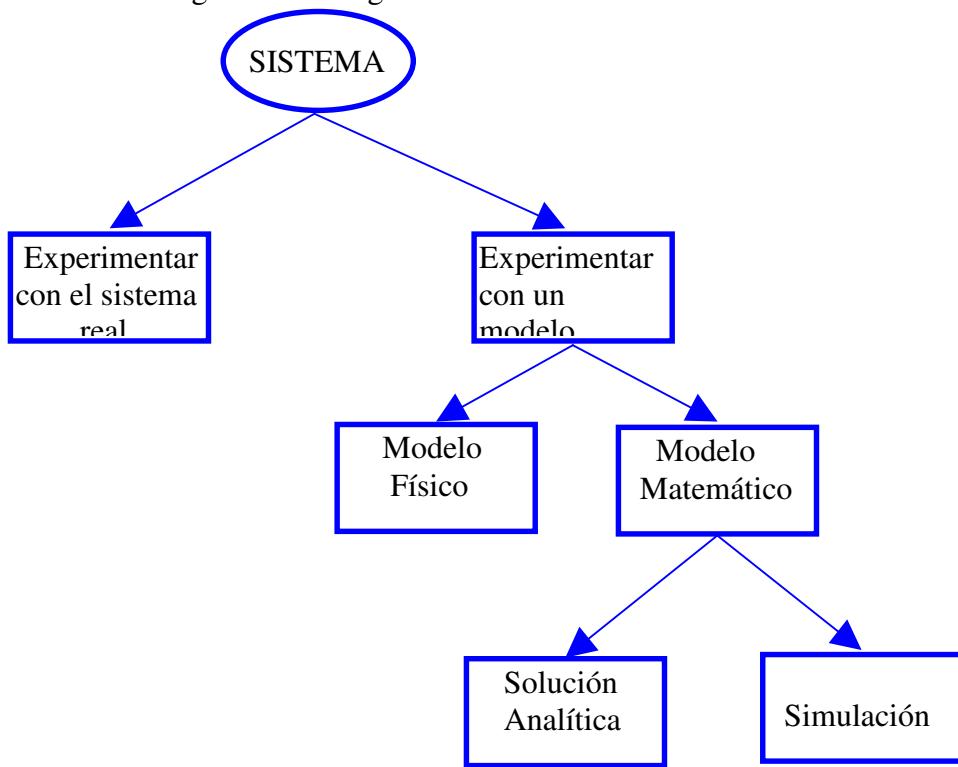


Figura 6. Formalización del proceso de decisión sobre el tipo de modelo según Law y Kelton

Al haberse hecho más fácil el uso de la simulación se ha incrementado notablemente su uso frente al de otros métodos para estudiar sistemas. De lo expuesto hasta aquí se desprende claramente que si bien la simulación tiene muchas ventajas, no deja de presentar algunos problemas, especialmente cuando se usa indebidamente, que cabe tener en cuenta para paliarlos o, si es posible, evitarlos, pues de lo contrario

pueden invalidar los resultados de un proyecto de simulación. Law y Kelton, [LAW], resumen en su texto la situación de la manera siguiente.

La simulación es recomendable, o incluso puede ser la única alternativa posible, para investigar sistemas complejos en los que estén presentes elementos estocásticos que difícilmente pueden ser tratados con la precisión adecuada en un modelo matemático. La simulación permite con relativa facilidad estimar el funcionamiento del sistema bajo condiciones de operación alternativas, o es la herramienta para comparar diseños alternativos de un sistema que tengan que satisfacer requerimientos específicos. La simulación permite mantener un mayor control sobre las condiciones experimentales que el que se puede mantener en la experimentación con el sistema físico. Por otra parte la simulación permite estudiar el comportamiento del sistema en períodos de tiempo de cualquier longitud, comprimidos a la duración de la ejecución del simulador en un computador.

Sin embargo, no se debe olvidar que la simulación de un modelo estocástico únicamente produce *estimaciones* de las características verdaderas del modelo para un conjunto particular de parámetros de entrada, lo que comporta la necesidad de diseñar adecuadamente los experimentos de simulación, y repetirlos en cantidad suficiente como para garantizar la calidad de las estimaciones. De ahí que los modelos de simulación no sean tan buenos cuando se pretende optimizar el rendimiento, como lo son cuando lo que se persigue es comparar entre sí diseños alternativos. En particular para sistemas para los que puede definirse un modelo analítico adecuado, que proporciona resultados exactos, este será preferible al modelo de simulación. No hay que olvidar en ningún caso que si un modelo no es una representación *válida* del sistema que se estudia la información que proporcione será de poca utilidad.

Modelos de simulación y modelos analíticos no deben considerarse siempre como distintos, en muchos casos pueden jugar un papel complementario sirviendo la simulación para verificar la validez de las hipótesis para el modelo analítico, o el modelo analítico para sugerir cuales son las alternativas razonables que hay que investigar por simulación.

Finalmente, suponiendo que se ha decidido proceder a un estudio de simulación, no hay que olvidar las recomendaciones metodológicas, que de no cumplirse pueden llevar al fracaso del estudio, como por ejemplo cuando no están bien definidos los objetivos del estudio, se construye un modelo con un nivel de detalle inadecuado, se concibe el estudio de simulación como una especie de ejercicio de programación de computadores, se utiliza inadecuadamente las capacidades de animación de resultados, no se identifican correctamente las fuentes de aleatoriedad del sistema, o se utilizan distribuciones de probabilidad que no se corresponden como es debido con dichas fuentes de aleatoriedad, o se utilizan métodos estadísticos incorrectos para analizar los resultados (por ejemplo los que suponen independencia), o se intenta extraer conclusiones de una única ejecución del simulador, etc.etc..

2. La Simulación de Sistemas Discretos: Lenguajes de Simulación de Sistemas Discretos

2.1 Metodología de la Construcción de Modelos de Simulación de Sistemas Discretos

Las ideas generales expuestas hasta aquí pueden resumirse en los términos siguientes, si las relaciones que constituyen el modelo matemático de un sistema lo permiten pueden obtenerse soluciones exactas de forma analítica que proporcionan respuestas a las cuestiones de interés sobre el sistema a través de las soluciones producidas por el modelo.

Cuando la complejidad de las ecuaciones del modelo hace inviables las soluciones analíticas hemos de recurrir a métodos numéricos para resolverlas. La Simulación se encuadra dentro de las técnicas numéricas a utilizar en estos casos.

Por *Simulación de Sistemas* entendemos el proceso de representar un sistema por medio de un modelo simbólico que puede manipularse por medio de un computador digital para que proporcione resultados numéricos. El computador permitirá evaluar numéricamente un modelo de simulación durante un intervalo de tiempo significativo para los objetivos del estudio, y recoger durante el mismo datos que posibiliten la estimación de las características deseadas del modelo.

Modelización y *Simulación* son pues los términos que utilizamos para designar el conjunto de actividades asociadas con la construcción de modelos de sistemas, y su utilización para realizar experimentos específicamente diseñados para estudiar aspectos concretos del comportamiento dinámico del sistema representado por el modelo. En particular el término *modelización* se refiere a las relaciones entre los sistemas reales y sus modelos, es decir el proceso a través del cual construimos el modelo que representa al sistema, mientras que *simulación* se refiere, sobre todo, a la manera en que el modelo es tratado en el computador.

La *Modelización por Simulación* (“Simulation Modelling”) supone que podemos describir un sistema en unos términos que son aceptables para un computador para poder describir sus cambios. Para aquellos sistemas cuyos cambios son discretos el concepto clave es el de la *descripción del estado del sistema*. Si un sistema puede ser caracterizado por un conjunto de variables, de manera que cada combinación de valores de las variables represente un *estado* único, o condición del sistema, entonces la manipulación de las variables *simula* el movimiento del sistema de estado a estado.

En esto es precisamente en lo que consiste la Simulación: *la representación de la dinámica del sistema moviéndolo de un estado a otro de acuerdo con reglas de operación bien definidas*.

En el caso de los sistemas discretos, la simulación se refiere a la modelización en un computador digital de un sistema cuyos cambios de estado tienen lugar de manera discreta como consecuencia de la ocurrencia de agentes del cambio de estado denominados sucesos.

Simular es, pues, duplicar el comportamiento dinámico de algún aspecto de un sistema, real o propuesto, sustituyendo las propiedades esenciales del sistema simulado por las propiedades del sistema que lo simula, el modelo.

Una vez construido, el modelo de simulación actúa como un laboratorio en el que se pueden verificar, comparar y ajustar fácilmente las diferentes alternativas de diseño. Mediante la ejecución del modelo en un ordenador se imitan las acciones del sistema propuesto, permitiendo extraer inferencias sobre las prestaciones del sistema global. Estas inferencias se realizan en base a las numerosas medidas de las prestaciones proporcionadas por la simulación, tales como las utilizaciones de la máquinas, los inventarios durante el proceso, los tiempos de espera de las piezas, y el "throughput" total, en el caso de la simulación de sistemas de producción, por ejemplo. Mediante este proceso de experimentación se puede elegir el mejor diseño del sistema.

La popularidad creciente de la Simulación es debida en parte a su fiel adherencia a los detalles. Las formulaciones analíticas, aunque ofrecen soluciones rápidas "closed-forms", tienden a ser restrictivas incluso para el más simple de los sistemas. Los detalles complicados se omiten con frecuencia para poder acomodar la formulación. La Simulación, por el contrario, puede proporcionar tantos detalles complicados como desee el analista. Todas las características relevantes del sistema pueden ser tenidas en cuenta, tales como la variabilidad de los tiempos de proceso de las operaciones, la fiabilidad de los equipos, "fixture restrictions", restricciones de almacenamiento intermedio durante los procesos, decisiones complejas de enrutamiento, políticas de operación y restricciones de programación de operaciones. Las hipótesis restrictivas de simplificación, tales como las exigencias de linealidad o de tiempos constantes, no son necesarias. El resultado es un contexto experimental más realista en el cual verificar y comprobar las estrategias de diseño alternativas, lo cual, a su vez, añade seguridad a las conclusiones a que se llega mediante este planteamiento.

En general las condiciones de utilización de la simulación son las siguientes según Shannon, [SHA]:

1. No existe una formulación matemática completa del problema, o no se han desarrollado aún los métodos analíticos para resolver el modelo matemático.
2. Existen los métodos analíticos, pero las hipótesis simplificadoras, necesarias para su aplicación, desvirtuan las soluciones obtenidas y su interpretación.
3. Los métodos analíticos existen, y en teoría están disponibles, pero los procedimientos numéricos son tan arduos y complejos que la simulación constituye un método más sencillo para obtener una solución.
4. Es deseable observar una historia simulada del proceso dentro de un horizonte temporal dado para poder estimar ciertos parámetros.
5. La simulación constituye la mejor alternativa por la dificultad de realizar experiencias en el contexto real.
6. Es necesario realizar una *compresión temporal* para estudiar la evolución del sistema a largo plazo.

Un modelo de simulación describe pues el *comportamiento dinámico* de un sistema en el tiempo. En la terminología que hemos introducido, el sistema está formado por **entidades**, elementos componentes cuyo comportamiento es seguido a través del sistema. Las entidades pueden pertenecer a diferentes clases, y poseen **atributos** identificativos que caracterizan su comportamiento, y por ende el del sistema. Las entidades de una clase dada tienen el mismo conjunto de atributos, aunque sus valores no tengan por qué ser idénticos, y tenderán a seguir pautas de comportamiento similares en el tiempo. Las entidades pueden intervenir en **actividades**, solas o conjuntamente con otras entidades.

En un instante determinado el sistema simulado tendrá una configuración dada, definida por las actividades de las entidades y los valores de sus atributos: *esta será la caracterización del estado del sistema*. El estado cambiará, discretamente, como consecuencia de la ocurrencia de un *sucedido*.

La secuencia de sucesos y los cambios de estado que determinan representan el comportamiento dinámico del sistema.

De este planteamiento conceptual de la simulación se desprende que el control del flujo del tiempo en el proceso de simulación es el aspecto clave del proceso de simulación. Tradicionalmente se ha propuesto dos procedimientos, un tratamiento síncrono y uno asíncrono. El tratamiento síncrono o de reparto del tiempo a intervalos regulares, consiste en desplazar el tiempo hacia adelante a intervalos iguales, es decir actualizar y examinar el modelo a intervalos regulares. Así, por ejemplo, si el intervalo considerado es de amplitud Δt , el modelo se actualizará en el instante $t+\Delta t$ para los cambios ocurridos en el intervalo $(t, t+\Delta t)$. Este enfoque plantea el problema de decidir cuál es la longitud apropiada de la amplitud del intervalo Δt . Evidentemente si la amplitud del intervalo de referencia es muy grande el comportamiento del modelo será muy burdo en relación con el del sistema puesto que será imposible tener en cuenta alguno de los cambios de estado que pueden ocurrir, mientras que por el contrario si el intervalo de tiempo es muy pequeño el modelo será reexaminado innecesariamente con gran frecuencia y el consiguiente incremento del coste computacional. Las dificultades para decidir apropiadamente en muchos casos la longitud del intervalo Δt hacen preferible utilizar un planteamiento asíncrono, en el que la amplitud del intervalo de tiempo es variable. En este caso el estado del modelo es examinado y actualizado únicamente cuando tiene lugar un cambio de estado. Esta es la situación típica cuando el cambio de estado es producido por la ocurrencia de un suceso, y puesto que entonces el tiempo es movido de suceso a suceso este enfoque recibe el nombre de técnica del suceso siguiente o dirigida por sucesos.

2.2 Características Generales de los Lenguajes de Simulación de Sistemas Discretos: La “Visión del Mundo” de un Lenguaje de Simulación.

La codificación de un modelo de simulación, de un sistema con sucesos discretos, en términos de un lenguaje de programación, de tipo general o especial para simulación, pone de manifiesto una serie de características comunes a todos ellos, como por ejemplo, Fishman [FIS] :

1. La generación de muestras de números aleatorios uniformemente distribuidos en $(0,1)$
2. La generación de muestras de variables aleatorias con distribuciones específicas.
3. Los mecanismos de control y flujo del tiempo durante la simulación.
4. La determinación del suceso que ha de ser tratado a continuación del que lo está siendo (técnica next-event).
5. La adición, supresión o actualización de registros en estructuras de datos ad hoc, como por ejemplo listas que representan secuencias de sucesos.
6. La recolección y el análisis de los datos generados por la simulación.
7. La elaboración de informes sobre los resultados obtenidos.
8. La detección de condiciones de error.

La existencia de estos y otros factores comunes a la mayor parte de los programas de simulación es lo que ha conducido al desarrollo de los lenguajes de simulación de propósito general, cuyo perfeccionamiento, estandarización y accesibilidad han sido los principales factores del incremento que ha experimentado el uso de la simulación en los últimos años. Una de las decisiones más importantes que ha de tomar un modelizador o analista para la realización de un estudio de simulación es la selección del lenguaje de computador en que va a programar el modelo. Algunas de las ventajas de

programar el modelo de simulación en un lenguaje de simulación en vez de hacerlo en un lenguaje general como FORTRAN, PASCAL, o C, son las siguientes:

1. Los lenguajes de simulación proporcionan automáticamente muchas, si no todas, las características necesarias para la programación de un modelo de simulación, lo que redunda en una reducción significativa del esfuerzo requerido para programar el modelo.
2. Proporcionan un marco de trabajo natural para el uso de modelos de simulación. Los bloques básicos de construcción del lenguaje son mucho más afines a los propósitos de la simulación que los de un lenguaje de tipo general.
3. Los modelos de simulación son mucho más fácilmente modificables cuando están escritos en un lenguaje de simulación, y por consiguiente se pueden ejecutar con mayor facilidad y flexibilidad los cambios requeridos por los experimentos de simulación.
4. Muchos de los lenguajes de simulación proporcionan una asignación dinámica de memoria durante la ejecución, cosa que no ocurre con todos los lenguajes de propósito general.
5. Facilitan una mejor detección de los errores, especialmente los inherentes a la lógica del proceso de simulación.

Los lenguajes de simulación para la simulación de sistemas discretos se eligen en base a lo atractivo de sus características, aunque hay una de ellas que resulta determinante e impone la naturaleza del lenguaje, es la **estrategia**, enfoque o **visión del mundo** inherente al lenguaje, utilizada para la selección del “suceso siguiente” y la gestión del tiempo. Kiviat, [KIV], en un trabajo seminal, clasificó los procedimientos de control de tiempos de los lenguajes de simulación de sucesos discretos, según la estrategia seguida en la selección de sucesos, en:

1. Event Scheduling
2. Activity Scanning
3. Process Interaction

Estrategias que conducen a la siguiente clasificación de los lenguajes de simulación, Fishman [FIS], Kiviat [KIV] :

ESTRATEGIA		
EVENT SCHEDULING	ACTIVITY SCANNING	PROCESS INTERACTION
GASP (II, IV) SIMSCRIPT II.5 SLAM SLAM II SIMAN	AS CSL ECSL ESP SIMON	GPSS (V, /H) Q-GERT SIMSCRIPT II.5 SLAM SLAM II SIMAN SIMULA

El hecho de que algunos lenguajes, SIMSCRIPT II.5, SLAM, SLAM II y SIMAN aparezcan tanto como lenguajes según una estrategia “event scheduling”, como “process interaction”, responde a que su estructura y naturaleza permiten al usuario emplear ambas estrategias, lo que refleja la tendencia creciente a la incorporación de múltiples visiones del mundo en la estructura de un solo lenguaje.

Con cualquiera de estas estrategias, cuando se selecciona el “suceso siguiente” para que sea procesado, se ejecuta la correspondiente “rutina de modelización” (bloque de código), para modelizar los cambios apropiados en el estado del modelo. *Los conceptos de suceso condicional, o incondicional, son básicos para las estrategias.* Un suceso incondicional es elegible para ser ejecutado cuando se llega al instante de tiempo para el que ha sido programado: *depende totalmente del tiempo.* Un suceso condicional *puede depender de condiciones adicionales distintas del tiempo.* Las condiciones se refieren usualmente a valores descriptores del estado de las componentes del sistema, por ejemplo, la desocupación, o el fin de la ocupación de un dispositivo del sistema modelizado.

Una característica común a los tres planteamientos es que los programas a que dan lugar tienen una estructura jerárquica de tres niveles, [FIS]:

- Nivel ejecutivo, o del programa de control
- Nivel de operaciones (secuencia de sentencias que constituyen el programa de computador que constituyen el modelo propiamente dicho)
- Nivel de rutinas de detalle, que ejecutan las acciones implicadas por cada una de las operaciones del modelo

El usuario de un lenguaje de simulación basado en uno de estos planteamientos no necesita conocer con detalle como está programado el ejecutivo, sino como funciona, es decir, lo que hemos llamado visión del mundo del lenguaje correspondiente. El ejecutivo es el que se encarga de identificar cuando tiene que ocurrir el siguiente suceso y que se ejecuten correctamente las operaciones por él implicadas en los momentos adecuados.

La estrategia **event scheduling** (ES) implica una secuencia de sucesos incondicionales a lo largo del tiempo. Por lo tanto, la **visión del mundo** de un lenguaje, o modelo expresado en tal lenguaje, que incorpore la estrategia ES consiste en considerar la operación del sistema como una secuencia temporal, ordenada, de sucesos incondicionales, que se implanta de la manera siguiente:

El procedimiento ES de control del tiempo selecciona de la lista de sucesos aquel cuyo tiempo de ocurrencia es el más próximo, resolviendo los empates por las prioridades asignadas o por defecto, actualiza el tiempo del reloj de simulación, igualando su valor al del instante en que ocurre el suceso, y llama a la rutina correspondiente al tratamiento del suceso.

Cualquier verificación de una condición diferente de la del tiempo del reloj debe realizarse dentro de las rutinas de tratamiento de los diferentes sucesos. Los sucesos se eligen y progresan sucesivamente hasta el instante de terminación. La figura 7, Emshoff [EMS], ilustra gráficamente el esquema de la rutina ejecutiva para la estrategia event scheduling.

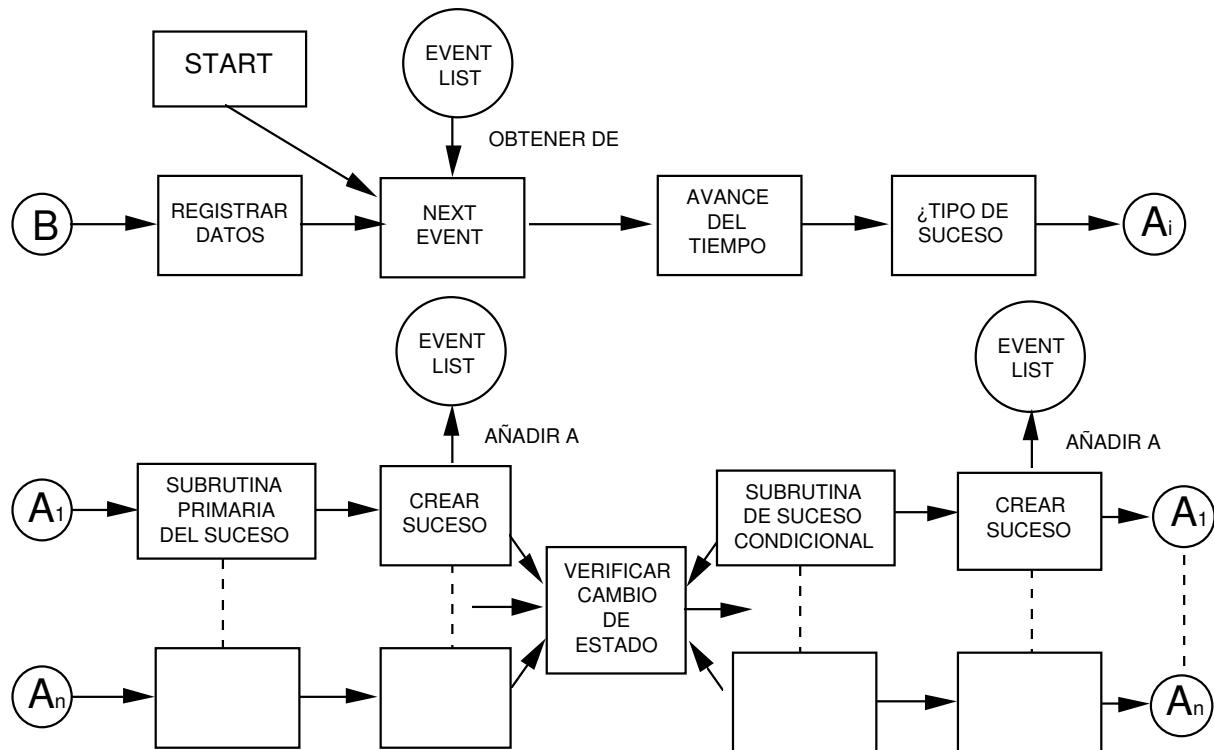


Figura 7. Estructura de la Estrategia ‘Event Scheduling’

La estrategia “**activity scanning**” (AS) elige el suceso siguiente basándose en ambos, el tiempo programado para la ocurrencia del suceso y la verificación de las condiciones. El concepto básico es la **actividad**, que conceptualmente es una transición de estado del sistema que requiere un periodo de tiempo. Una actividad se representa habitualmente mediante dos sucesos distintos que marcan respectivamente el principio y el final de la misma.

La visión del mundo que corresponde a un lenguaje, o modelo de simulación programado en un lenguaje, basado en la estrategia AS, considera que el sistema consta de componentes que desarrollan actividades sujetas a actividades específicas.

En la implantación de esta estrategia se particularizan las componentes de “**tipo_activo**”, o componentes “móviles” del sistema, que pueden tener asociada una rutina de actividad que cuando se ejecuta modela las fases de la componente; cada componente de “**tipo_activo**” tiene también asociado un reloj interno que indica en qué instante la componente vuelve a poder ser tenida en cuenta para ser activada de nuevo, y una rutina de condición para determinar si se cumplen las condiciones, distintas del tiempo, requeridas para su reactivación. *El procedimiento de control del tiempo AS explora las actividades en orden de prioridad para su elegibilidad en el tiempo y otras condiciones*

complementarias de activación, y ejecuta la rutina correspondiente a la actividad de la primera componente cuyas condiciones de activación se satisfacen.

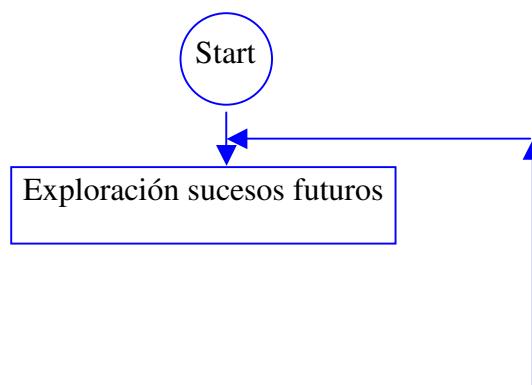
La estrategia “**Process Interaction**” (PI) tiene características híbridas que la relacionan tanto con la estrategia ES como con la AS

La visión del mundo de un lenguaje, o modelo de simulación programado en un lenguaje, basado en una estrategia PI contempla el progreso de las componentes del sistema a través de una secuencia de pasos, o procesos, cada uno de los cuales puede tener dos posibles componentes, un segmento de condición, cuya ejecución identifica si se puede pasar a ejecutar la segunda componente, un segmento de acción.

La implantación de la estrategia PI en un lenguaje como el GPSS, (General Purpose Simulation System, lenguaje pionero desarrollado por Gordon en 1964, Gordon [GRE], que a pesar de los años transcurridos mantiene su vigencia gracias a las continuas actualizaciones, Schriber [SCH]) se basa en un procedimiento de control del tiempo que utiliza dos listas de sucesos:

- Una lista de *sucesos futuros* (Future Events List, FEL), que contiene información sobre los eventos programados para ser ejecutados en un instante posterior.
- Una lista de *sucesos en curso* (Current Events List, CEL), que contiene información sobre los eventos que ya son elegibles desde el punto de vista del momento en que tienen que ser ejecutados, pero cuyas otras condiciones pueden no satisfacerse aún.

La información sobre cada suceso contiene una indicación sobre la localización del paso del proceso en curso en que se encuentra la componente. Con el avance del tiempo, (exploración de FEL para determinar en qué instante ocurrirá el siguiente suceso, y avanzar el reloj de simulación hasta ese instante) todos los sucesos programados para ese instante pasan de FEL a CEL, (movimientos entre listas, aquellas entidades de FEL cuya reactivación ha de tener lugar en ese instante del tiempo marcado por el reloj de simulación pasan de FEL a CEL) entonces tiene lugar una exploración de CEL, consistente en la evaluación de las condiciones de la rutina de entrada para determinar si la correspondiente componente puede proceder al paso siguiente, en caso afirmativo se ejecuta el segmento de acción del paso. Una componente se desplaza a lo largo de tantos pasos sucesivos del modelo como le sea posible, por ejemplo, en tanto en cuanto el tiempo no tenga que avanzar y sea afirmativa la respuesta a la evaluación de los segmentos de condición. Cuando una componente se “para”, debido al tiempo o a otras condiciones, entonces la exploración pasa a tener en cuenta el elemento siguiente de CEL; *cuando no se puede continuar moviendo ninguna de las componentes de CEL entonces se produce el siguiente avance en el flujo del tiempo*. La figura 9 esquematiza la estructura del ejecutivo para la estrategia de interacción de procesos.



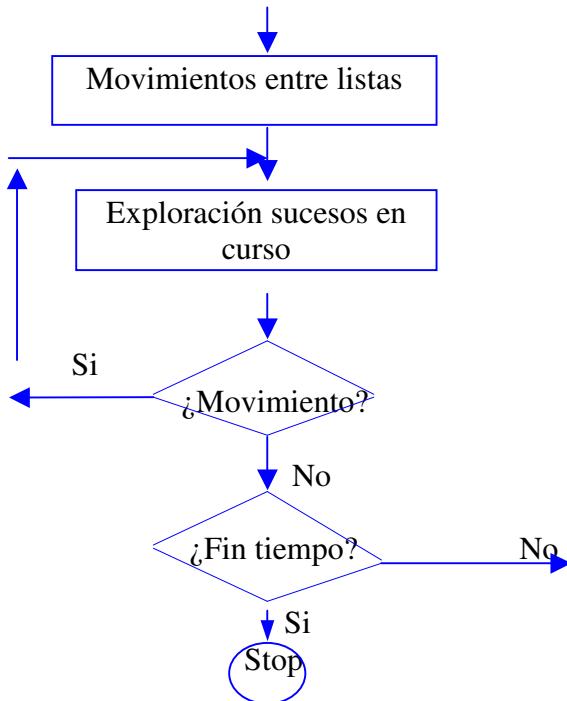


Figura 9: Ejecutivo de la estrategia de interacción de procesos

Si un lenguaje de simulación incluye más de una visión del mundo o estrategia, lo que cada vez es más corriente en el diseño de nuevos lenguajes o nuevas versiones de antiguos, el modelizador está aún constreñido por la estrategia elegida para representar una porción de su modelo, *de ahí la necesidad de una buena comprensión de los conceptos subyacentes a cada estrategia y el condicionamiento que imponen sobre la representación del modelo en sus términos*.

2.3 Análisis Algorítmico de las Estrategias de Simulación de Sistemas Discretos

Cada una de las estrategias de simulación que hemos propuesto son susceptibles de una representación algorítmica genérica, así como las rutinas para el tratamiento de sucesos de la estrategia ES, las rutinas de condición y las de actividades para la estrategia AS, y los segmentos de condición y acción para la estrategia PI.

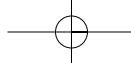
En la representación algorítmica que vamos a dar se supone que los algoritmos utilizan algún conjunto de procedimientos de gestión de estructuras de datos (por ejemplo listas), para ayudar a la introducción y recuperación de las representaciones de la información sobre los sucesos.

En el planteamiento que vamos a proponer, y para mantener el carácter genérico de la representación, evitaremos suposiciones sobre las estructuras de datos que se pueden utilizar en las implantaciones específicas de las estrategias. La descripción algorítmica que vamos a hacer será, pues, precisa solo en la descripción de los aspectos esenciales de la gestión del tiempo, incluyendo las implicaciones correspondientes a las rutinas de modelización, aunque excluyendo los detalles que no dependen de la estrategia.

El punto de partida para esta representación algorítmica es un análisis de las características que dependen de la estrategia adoptada y de las que dependen del problema específico que se simula. De acuerdo con Kiviat [KIV], los lenguajes de simulación ofrecen, en general, las siguientes capacidades y características:

- 1.** Modelización del estado del sistema.
- 2.** Modelización de la dinámica del sistema.
- 3.** Herramientas para el muestreo estadístico.
- 4.** Especificaciones sobre el proceso de recogida de datos.
- 5.** Recogida de datos generados por la simulación.
- 6.** Análisis de los datos recogidos.
- 7.** Presentación de resultados
- 8.** Monitorización de la simulación y herramientas de depuración del modelo.
- 9.** Procedimientos de inicialización de la simulación.
- 10.** Herramientas auxiliares complementarias.

Las características 1, 2 y 9 son intrínsecas, dependen de la estrategia elegida y por lo tanto son las que pueden ser objeto de una descripción algorítmica genérica de la estrategia, mientras que las restantes son relativamente independientes de la estrategia. A partir de la identificación de los elementos intrínsecos podemos dar las siguientes descripciones de las estrategia en pseudocódigo, Hooper [HOO]:



Chapter 5

Causation and Experimental Design

Causal Explanation

What Causes What?

Association
Time Order
Nonspuriousness
Mechanism
Context

Why Experiment?

What If a True Experiment Isn't Possible?

Nonequivalent Control Group Designs
Before-and-After Designs
Ex Post Facto Control Group Designs

What Are the Threats to Validity in Experiments?

Threats to Internal Causal Validity

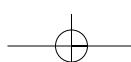
Noncomparable Groups
Endogenous Change
History
Contamination
Treatment Misidentification
Generalizability
Sample Generalizability
Cross-Population
Generalizability
Interaction of Testing and Treatment

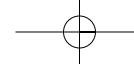
How Do Experimenters Protect Their Subjects?

Deception
Selective Distribution of Benefits

Conclusion

Identifying causes—figuring out why things happen—is the goal of most social science research. Unfortunately, valid explanations of the causes of social phenomena do not come easily. Why did the homicide rate in the United States drop for 15 years and then start to rise in 1999 (Butterfield, 2000:12)? Was it because of changes in the style of policing (Radin, 1997:B7) or because of changing attitudes among young people (Butterfield, 1996a)? Was it due to variation in





patterns of drug use (Krauss, 1996) or to tougher prison sentences (Butterfield, 1996a) or to more stringent handgun regulations (Butterfield, 1996b)? Did better emergency medical procedures result in higher survival rates for victims (Ramirez, 2002)? If we are to evaluate these alternative explanations we must design our research strategies carefully.

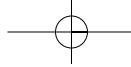
This chapter considers the meaning of causation, the criteria for achieving causally valid explanations, the ways in which experimental and quasi-experimental research designs seek to meet these criteria, and the difficulties that can sometimes result in invalid conclusions. By the end of the chapter, you should have a good grasp of the meaning of causation and the logic of experimental design. Most social research, both academic and applied, uses data collection methods other than experiments. But because experimental designs are the best way to evaluate causal hypotheses, a better understanding of them will help you to be aware of the strengths and weaknesses of other research designs that we will consider in subsequent chapters.

CAUSAL EXPLANATION

A cause is an explanation for some characteristic, attitude, or behavior of groups, individuals, or other entities (such as families, organizations, or cities) or for events. For example, Sherman and Berk (1984) conducted a study to determine whether adults who were accused of a domestic violence offense would be less likely to repeat the offense if police arrested them rather than just warned them. Their conclusion that this hypothesis was correct meant that they believed police response had a **causal effect** on the likelihood of committing another domestic violence offense.

Causal effect: The finding that change in one variable leads to change in another variable, *ceteris paribus* (other things being equal). *Example:* Individuals arrested for domestic assault tend to commit fewer subsequent assaults than similar individuals who are accused in the same circumstances but are not arrested.

More specifically, a causal effect is said to occur if variation in the independent variable is followed by variation in the dependent variable, when all other things are equal (*ceteris paribus*). For instance, we know that for the most part men earn more income than women do. But is this because they are men—or could it be due to higher levels of education, or to longer tenure in their jobs (with no pregnancy breaks), or is it the kinds of jobs men go into as compared to those that women choose? We want to know if men earn more than women, *ceteris paribus*—other things (job, tenure, education, etc.) being equal.



We admit that you can legitimately argue that “all” other things can’t literally be equal: We can’t compare the same people at the same time in exactly the same circumstances except for the variation in the independent variable (King, Keohane, & Verba, 1994). However, you will see that we can design research to create conditions that are very comparable so that we can isolate the impact of the independent variable on the dependent variable.

WHAT CAUSES WHAT?

Five criteria should be considered in trying to establish a causal relationship. The first three criteria are generally considered as requirements for identifying a causal effect: (1) empirical association, (2) temporal priority of the independent variable, and (3) nonspuriousness. You must establish these three to claim a causal relationship. Evidence that meets the other two criteria—(4) identifying a causal mechanism, and (5) specifying the context in which the effect occurs—can considerably strengthen causal explanations.

Research designs that allow us to establish these criteria require careful planning, implementation, and analysis. Many times, researchers have to leave one or more of the criteria unmet and are left with some important doubts about the validity of their causal conclusions, or they may even avoid making any causal assertions.

Association

The first criterion for establishing a causal effect is an empirical (or observed) **association** (sometimes called a *correlation*) between the independent and dependent variables. They must vary together so when one goes up (or down), the other goes up (or down) at the same time. For example: When cigarette smoking goes up, so does lung cancer. The longer you stay in school, the more money you will make later in life. Single women are more likely to live in poverty than married women. When income goes up, so does overall health. In all of these cases, a change in an independent variable correlates, or is associated with, a change in a dependent variable. If there is no association, there cannot be a causal relationship. For instance, empirically there seems to be no correlation between the use of the death penalty and a reduction in the rate of serious crime. That may seem unlikely to you, but empirically it is the case: There is no correlation. So there cannot be a causal relationship.

Time Order

Association is necessary for establishing a causal effect, but it is not sufficient. We must also ensure that the variation in the independent variable came before variation in the dependent variable—the cause must come before its presumed



effect. This is the criterion of **time order**, or the temporal priority of the independent variable. Motivational speakers sometimes say that to achieve success (the dependent variable in our terms), you need to really believe in yourself (the independent variable). And it is true that many very successful politicians, actors, and businesspeople seem remarkably confident—there is an association. But it may well be that their confidence is the result of their success, not its cause. Until you know which came first, you can't establish a causal connection.

Nonspuriousness

The third criterion for establishing a causal effect is **nonspuriousness**. *Spurious* means false or not genuine. We say that a relationship between two variables is **spurious** when it is actually due to changes in a third variable, so what appears to be a direct connection is in fact not one. Have you heard the old adage “Correlation does not prove causation”? It is meant to remind us that an association between two variables might be caused by something else. If we measure children’s shoe sizes and their academic knowledge, for example, we will find a positive association. However, the association results from the fact that older children have larger feet as well as more academic knowledge; a third variable (age) is affecting both shoe size and knowledge, so that they correlate. But one doesn’t cause the other. Shoe size does not cause knowledge, or vice versa. The association between the two is, we say, spurious.

If this point seems obvious, consider a social science example. Do schools with better resources produce better students? There is certainly a correlation, but consider the fact that parents with more education and higher income tend to live in neighborhoods that spend more on their schools. These parents are also more likely to have books in the home and to provide other advantages for their children (see Exhibit 5.1). Maybe parents’ income causes variation in both school resources and student performance. If so, there would be an association between school resources and student performance, but it would be at least partially spurious. What we want, then, is *nonspuriousness*.

Mechanism

A causal **mechanism** is the process that creates the connection between the variation in an independent variable and the variation in the dependent variable that it is hypothesized to cause (Cook & Campbell, 1979:35; Marini & Singer, 1988). Many social scientists (and scientists in other fields) argue that no causal explanation is adequate until a mechanism is identified.

For instance, there seems to be an empirical association at the individual level between poverty and delinquency: Children who live in impoverished homes seem

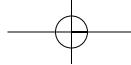
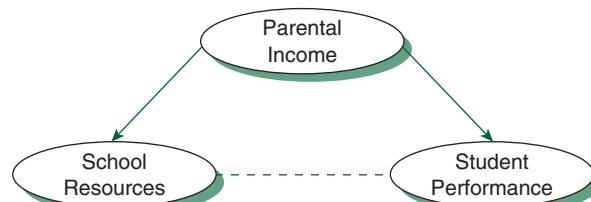


Exhibit 5.1 A Spurious Relationship Revealed

School resources are associated with student performance; apparently, a causal relation



But in fact, parental income (a third variable) influences both school resources and student performance, creating the association

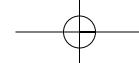


more likely to be involved in petty crime. But why? Some researchers have argued for a *mechanism* of low parent/child attachment, inadequate supervision of children, and erratic discipline as the means by which poverty and delinquency are connected (Sampson & Laub, 1994). In this way, figuring out some aspects of the process by which the independent variable influenced the variation in the dependent variable can increase confidence in our conclusion that there was a causal effect (Costner, 1989).

Context

No cause has its effect apart from some larger **context** involving other variables. When, for whom, and in what conditions does this effect occur? A cause is really one among a set of interrelated factors required for the effect (Hage & Meeker, 1988; Papineau, 1978). Identification of the context in which a causal effect occurs is not itself a criterion for a valid causal conclusion, and it is not always attempted; but it does help us to understand the causal relationship.

You may hypothesize, for example, that if you offer employees higher wages to work harder, they will indeed work harder; and in the context of America, this seems to indeed be the case. Incentive pay causes harder work. But in noncapitalist societies, workers often want only enough money to meet their basic needs and would rather work less than drive themselves hard just to have more money.



In America, the correlation of incentive pay with greater effort seems to work; in medieval Europe, for instance, it did not (Weber, 1992).

As another example, in America in the 1960s, children of divorced parents (“from a broken home”) were more likely to suffer from a variety of problems; they lived in a context of mostly intact families. In 2006, many parents are divorced, and the causal link between divorced parents and social pathology no longer seems to hold (Coontz, 1997).

WHY EXPERIMENT?

Experimental research provides the most powerful design for testing causal hypotheses because it allows us to confidently establish the first three criteria for causality—association, time order, and nonspuriousness. **True experiments** have at least three features that help us meet these criteria:

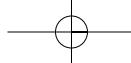
1. Two comparison groups (in the simplest case, an experimental group and a control group), to establish association
2. Variation in the independent variable before assessment of change in the dependent variable, to establish time order
3. Random assignment to the two (or more) comparison groups, to establish nonspuriousness

We can determine whether an association exists between the independent and dependent variables in a true experiment because two or more groups differ in terms of their value on the independent variable. One group receives some “treatment” that is a manipulation of the value of the independent variable. This group is termed the **experimental group**. In a simple experiment, there may be one other group that does not receive the treatment; it is termed the **control group**.

Experimental group: In an experiment, the group of subjects that receives the treatment or experimental manipulation.

Control group: A comparison group that receives no treatment.

Consider an example in detail (see the simple diagram in Exhibit 5.2). Does drinking coffee improve one’s writing of an essay? Imagine a simple experiment. Suppose you believe that drinking two cups of strong coffee before class will help you in writing an in-class essay. But other people think that coffee makes them too nervous and “wired” and so doesn’t help in writing the essay. To test your hypothesis (“coffee drinking causes improved performance”), you need to


Exhibit 5.2 A True Experiment

Experimental Group: R O₁ X O₂

Comparison Group: R O₁ O₂

Key: R = Random assignment
 O = Observation (pretest [O₁] or posttest [O₂])
 X = Experimental treatment

Experimental Group	O ₁ Pretest Essay	X Coffee	O ₂ Posttest Essay
Comparison Group	Pretest Essay		Posttest Essay

compare two groups of subjects, a control group and an experimental group. First, the two groups will sit and write an in-class essay. Then, the control group will drink no coffee while the experimental group will drink two cups of strong coffee. Next, both groups will sit and write another in-class essay. At the end, all of the essays will be graded and you will see which group improved more. Thus, you may establish *association*.

You may find an association outside the experimental setting, of course, but it won't establish time order. Perhaps good writers hang out in cafés and coffee houses, and then start drinking lots of coffee. So there would be an association, but not the causal relation we're looking for. By controlling who gets the coffee, and when, we establish *time order*.

All true experiments have a **posttest**—that is, a measurement of the outcome in both groups after the experimental group has received the treatment. In our example, you grade the papers. Many true experiments also have **pretests** that measure the dependent variable before the experimental intervention. A pretest is exactly the same as a posttest, just administered at a different time. Strictly speaking, though, a true experiment does not require a pretest. When researchers use random assignment, the groups' initial scores on the dependent variable and on all other variables are very likely to be similar. Any difference in outcome between the experimental and comparison groups is therefore likely to be due to the intervention (or to other processes occurring during the experiment), and the likelihood of a difference just on the basis of chance can be calculated.



Finally, it is crucial that the two groups be more or less equal at the beginning of the study. If you let students choose which group to be in, the more ambitious students may pick the coffee group, hoping to stay awake and do better on the paper. Or people who simply don't like the taste of coffee may choose the non-coffee group. Either way, your two groups won't be equivalent at the beginning of the study, and so any difference in their writing may be the result of that initial difference (a source of spuriousness), not the drinking of coffee.

So you randomly sort the students into the two different groups. You can do this by flipping a coin for each one of them, or by pulling names out of a hat, or by using a random number table as described in the previous chapter. In any case, the subjects themselves should not be free to choose, nor should you (the experimenter) be free to put them into whatever group you want. (If you did that, you might unconsciously put the better students into the coffee group, hoping to get the results you're looking for.) Thus we hope to achieve nonspuriousness.

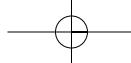
Note that the random assignment of subjects to experimental and comparison groups is not the same as random sampling of individuals from some larger population (see Exhibit 5.3). In fact, **random assignment (randomization)** does not help at all to ensure that the research subjects are representative of some larger population; instead, representativeness is the goal of random sampling. What random assignment does—create two (or more) equivalent groups—is useful for ensuring internal validity, not generalizability.

Matching is another procedure sometimes used to equate experimental and comparison groups, but by itself it is a poor substitute for randomization. Matching of individuals in a treatment group with those in a comparison group might involve pairing persons on the basis of similarity of gender, age, year in school, or some other characteristic. The basic problem is that, as a practical matter, individuals can be matched on only a few characteristics; unmatched differences between the experimental and comparison groups may still influence outcomes.

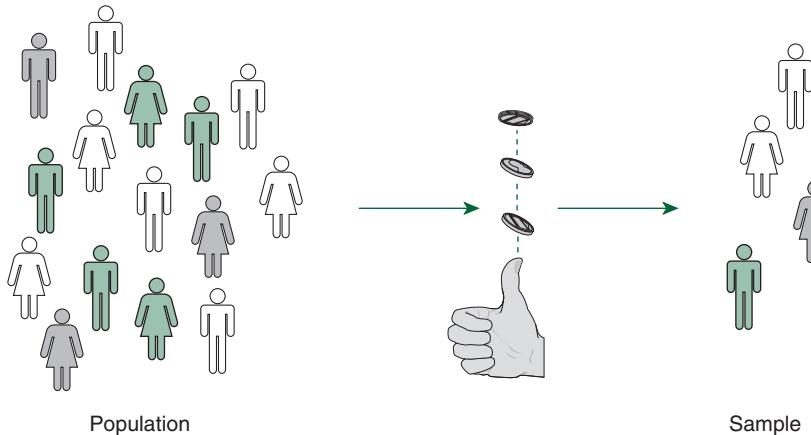
These defining features of true experimental designs give us a great deal of confidence that we can meet the three basic criteria for identifying causes: association, time order, and nonspuriousness. However, we can strengthen our understanding of causal connections, and increase the likelihood of drawing causally valid conclusions, by also investigating causal mechanism and causal context.

WHAT IF A TRUE EXPERIMENT ISN'T POSSIBLE?

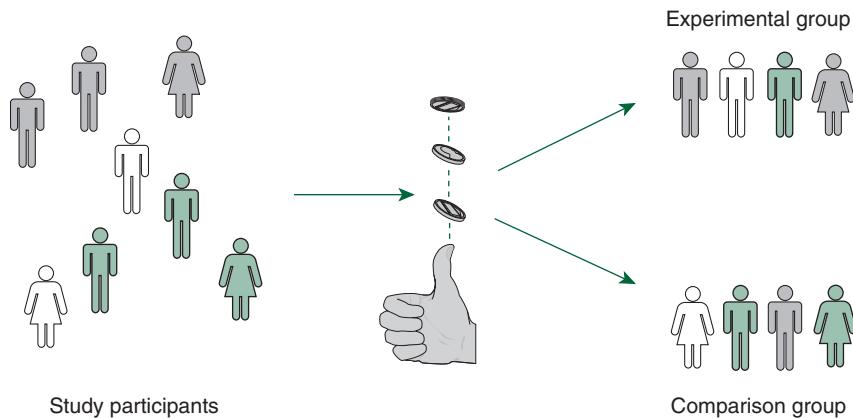
Often, testing a hypothesis with a true experimental design is not feasible. A true experiment may be too costly or take too long to carry out; it may not be ethical to randomly assign subjects to the different conditions; or it may be too late to do


Exhibit 5.3 Random Sampling Versus Random Assignment

Random sampling (a tool for ensuring generalizability):
 Individuals are randomly selected from a population to participate in a study.

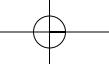


Random assignment, or randomization (a tool for ensuring internal validity):
 Individuals who are to participate in a study are randomly divided into an experimental group and a comparison group.



so. Researchers may instead use “quasi-experimental” designs that retain several components of experimental design but differ in important details.

In **quasi-experimental design**, a comparison group is predetermined to be comparable to the treatment group in critical ways, such as being eligible for the same services or being in the same school cohort (Rossi & Freeman, 1989:313). These research designs are only “quasi”-experimental because subjects are not



randomly assigned to the comparison and experimental groups. As a result, we cannot be as confident in the comparability of the groups as in true experimental designs. Nonetheless, in order to term a research design quasi-experimental, we have to be sure that the comparison groups meet specific criteria.

We will discuss here the two major types of quasi-experimental designs, as well as one type—*ex post facto* (after the fact) control group design—that is often mistakenly termed quasi-experimental (other types can be found in Cook & Campbell, 1979, and Mohr, 1992):

- *Nonequivalent control group designs*—**Nonequivalent control group designs** have experimental and comparison groups that are designated before the treatment occurs but are not created by random assignment.
- *Before-and-after designs*—**Before-and-after designs** have a pretest and posttest but no comparison group. In other words, the subjects exposed to the treatment serve, at an earlier time, as their own control group.
- *Ex post facto control group designs*—These designs use nonrandomized control groups designated after the fact.

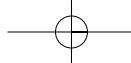
Exhibit 5.4 diagrams one study using the *ex post facto* control group design and another study using the multiple group before-and-after design, one type of before-and-after design. (The diagram for an *ex post facto* control group design is the same as for a nonequivalent control group design, but the two types of experiment differ in how people are able to join the groups.)

If quasi-experimental designs are longitudinal, they can establish time order. Where these designs are weaker than true experiments is in establishing the nonspuriousness of an observed association—that it does not result from the influence of some third, uncontrolled variable. On the other hand, because these quasi-experiments do not require the high degree of control necessary in order to achieve random assignment, quasi-experimental designs can be conducted using more natural procedures in more natural settings, so we may be able to achieve a more complete understanding of causal context. In identifying the mechanism of a causal effect, though, quasi-experiments are neither better nor worse than experiments.

Nonequivalent Control Group Designs

In this type of quasi-experimental design, a comparison group is selected so as to be as comparable as possible to the treatment group. Two selection methods can be used:

Individual matching—Individual cases in the treatment group are matched with similar individuals in the comparison group. This can sometimes create a comparison


Exhibit 5.4 Quasi-Experimental Designs
Nonequivalent control group design:

Experimental group:	O ₁	X _a	O ₂
Comparison group 1:	O ₁	X _b	O ₂
Comparison group 2:	O ₁	X _c	O ₂
	<i>Pretest</i>	<i>Treatment</i>	<i>Posttest</i>
<i>Team Interdependence</i>	Group	Team performance	Independent tasks
	Hybrid	Team performance	Mixed tasks
	Individual	Team performance	Individual tasks
			Team performance

Before-and-after design:
Soap-opera suicide and actual suicide (Phillips, 1982)

Experimental group:	O ₁₁	X ₁	O ₂₁
	O ₁₂	X ₂	O ₂₂
	O ₁₃	X ₃	O ₂₃
	O ₁₄	X ₄	O ₂₄
	<i>Pretest</i>	<i>Treatment</i>	<i>Posttest</i>
	Suicide rate	Soap-opera suicides	Suicide rate

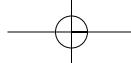
Key: O = Observation (pretest or posttest)

X = Experimental treatment

Source: Ruth Wageman, 1995. "Interdependence and Group Effectiveness." *Administrative Science Quarterly*, 40:145–180. Reprinted with permission.

group that is very similar to the experimental group, such as when Head Start participants were matched with their siblings to estimate the effect of participation in Head Start. However, in many studies it may not be possible to match on the most important variables.

Aggregate matching—In most situations when random assignment is not possible, the second method of matching makes more sense: identifying a comparison group that matches the treatment group in the aggregate rather than trying to match individual cases. This means finding a comparison group that has similar distributions on key variables: the same average age, the same percentage female, and so on. For this design to be considered quasi-experimental, however, it is important that individuals must themselves have chosen to be in the treatment group or the control group.



Nonequivalent control group designs allow you to determine whether an association exists between the presumed cause and effect.

Before-and-After Designs

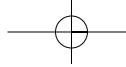
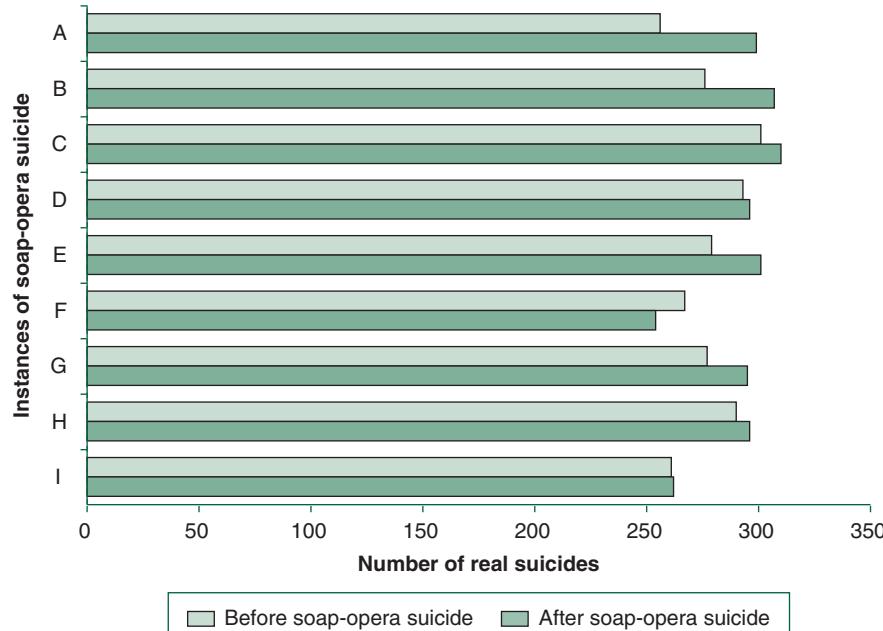
The common feature of before-and-after designs is the absence of a comparison group: All cases are exposed to the experimental treatment. The basis for comparison is instead provided by the pretreatment measures in the experimental group. These designs are thus useful for studies of interventions that are experienced by virtually every case in some population, such as total coverage programs like Social Security or single-organization studies of the effect of a new management strategy.

The simplest type of before-and-after design is the fixed-sample panel design. As you may recall from Chapter 2, in a panel design the same individuals are studied over time, the research may entail one pretest and one posttest. However, this type of before-and-after design does not qualify as a quasi-experimental design because comparing subjects to themselves at just one earlier point in time does not provide an adequate comparison group. Many influences other than the experimental treatment may affect a subject following the pretest—for instance, basic life experiences for a young subject.

David P. Phillips's (1982) study of the effect of TV soap-opera suicides on the number of actual suicides in the United States illustrates a more powerful **multiple group before-and-after design**. In this design, before-and-after comparisons are made of the same variables between different groups. Phillips identified 13 soap-opera suicides in 1977 and then recorded the U.S. suicide rate in the weeks prior to and following each TV story. In effect, the researcher had 13 different before-and-after studies, one for each suicide story. In 12 of these 13 comparisons, deaths due to suicide increased from the week before each soap-opera suicide to the week after (see Exhibit 5.5). Phillips also found similar increases in motor-vehicle deaths and crashes during the same period, some portion of which reflects covert suicide attempts. (Despite his clever design, however, some prominent researchers have disputed his findings.)

Another type of before-and-after design involves multiple pretest and posttest observations of the same group. **Repeated measures panel designs** include several pretest and posttest observations, allowing the researcher to study the process by which an intervention or treatment has an impact over time; hence, they are better than a simple before-and-after study.

Time series designs include many (preferably 30 or more) such observations in both pretest and posttest periods. They are particularly useful for studying the impact of new laws or social programs that affect large numbers of people and that are readily assessed by some ongoing measurement. For example, we might use a time series design to study the impact of a new seat-belt law on the severity of injuries in


Exhibit 5.5 Real Suicides and Soap-Opera Suicides


Source: David P. Phillips, 1982. "The Impact of Fictional Television Stories on U.S. Adult Fatalities: New Evidence on the Effect of the Mass Media on Violence." *American Journal of Sociology*, 87 (May 1982):1340. Copyright © 1982 by the University of Chicago Press. Reprinted with permission.

automobile accidents, using a monthly state government report on insurance claims. Special statistics are required to analyze time series data, but the basic idea is simple: identify a trend in the dependent variable up to the date of the intervention, then project the trend into the postintervention period. This *projected* trend is then compared to the *actual* trend of the dependent variable after the intervention. A substantial disparity between the actual and projected trends is evidence that the intervention or event had an impact (Rossi & Freeman, 1989:260–261, 358–363).

How well do these before-and-after designs meet the five criteria for establishing causality? The before-after comparison enables us to determine whether an *association* exists between the intervention and the dependent variable (because we can determine whether there was a change after the intervention). They also clarify whether the change in the dependent variable occurred after the intervention, so *time order* is not a problem. However, there is no control group so we cannot rule out the influence of extraneous factors as the actual cause of the change we observe; *spuriousness* may be a problem. Some other event may have occurred during the



study that resulted in a change in posttest scores. Overall, the longitudinal nature of before-and-after designs can help to identify causal mechanisms, while the loosening of randomization requirements makes it easier to conduct studies in natural settings, where we learn about the influence of contextual factors.

Ex Post Facto Control Group Designs

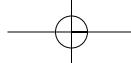
The **ex post facto control group design** appears to be very similar to the nonequivalent control group design and is often confused with it, but it does not meet as well the criteria for quasi-experimental designs. Like nonequivalent control group designs, this design has experimental and comparison groups that are not created by random assignment. But unlike the groups in nonequivalent control group designs, the groups in ex post facto designs are designated after the treatment has occurred. The problem with this is that if the treatment takes any time at all, people with particular characteristics may select themselves for the treatment or avoid it. Of course, this makes it difficult to determine whether an association between group membership and outcome is spurious. However, the particulars will vary from study to study; in some circumstances we may conclude that the treatment and control groups are so similar that causal effects can be tested (Rossi & Freeman, 1989:343–344).

Susan Cohen and Gerald Ledford's (1994) study of the effectiveness of self-managing teams used a well-constructed ex post facto design. They studied a telecommunications company with some work teams that were self-managing and some that were traditionally managed (meaning that a manager was responsible for the team's decisions). Cohen and Ledford found the self-reported quality of work life to be higher in the self-managing groups than in the traditionally managed groups.

WHAT ARE THE THREATS TO VALIDITY IN EXPERIMENTS?

Experimental designs, like any research design, must be evaluated for their ability to yield valid conclusions. Remember, there are three kinds of validity: internal (or causal), external (or generalizability), and measurement. True experiments are good at producing internal validity, but they fare less well in achieving external validity (generalizability). Quasi-experiments may provide more generalizable results than true experiments but are more prone to problems of internal invalidity. Measurement validity is also a central concern for both kinds of research, but even true experimental design offers no special advantages or disadvantages in measurement.

In general, nonexperimental designs such as those used in survey research and field research offer less certainty of internal validity, a greater likelihood of generalizability, and no particular advantage or disadvantage in terms of measurement

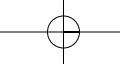

Exhibit 5.6 Threats to Internal Validity

Problem	Example	Type
Selection	Girls who choose to see a therapist are not representative of population.	Noncomparable Groups
Mortality	Students who most dislike college drop out, so aren't surveyed.	Noncomparable Groups
Instrument Decay	Interviewer tires losing interest in later interviews, so poor answers.	Noncomparable Groups
Testing	If someone has taken the SAT before, they are familiar with the format, so do better.	Endogenous Change
Maturation	Everyone gets older in high school; it's not the school's doing.	Endogenous Change
Regression	The lowest-ranking students on IQ must improve their rank; they can't do worse.	Endogenous Change
History	The O.J. Simpson trial affects members of diversity workshops.	History
Contamination	“John Henry” effect; people in study compete with one another.	Contamination
Experimenter Expectation	Researchers unconsciously help their subjects, distorting results.	Treatment Misidentification
Placebo Effect	Fake pills in medical studies produce improved health.	Treatment Misidentification
Hawthorne Effect	Workers enjoy being subjects and work harder.	Treatment Misidentification

validity. We will introduce survey and field research designs in the following chapters; in this section we focus on the ways in which experiments help (or don't help) to resolve potential problems of internal validity and generalizability.

Threats to Internal Causal Validity

The following sections discuss 10 threats to validity (also referred to as “sources of invalidity”) that occur frequently in social science research (see Exhibit 5.7). These “threats” exemplify five major types of problems that arise in research design.



Noncomparable Groups

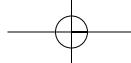
The problem of noncomparable groups occurs when the experimental group and the control group are not really comparable—that is, when something interferes with the two groups being essentially the same at the beginning (or end) of a study.

- *Selection bias*—Occurs when the subjects in your groups are initially different. If the ambitious students decide to be in the “coffee” group, you’ll think their performance was helped by coffee—but it could have been their ambition.

Everyday examples of selection bias are everywhere. Harvard graduates are very successful people; but Harvard *admits* students who are likely to be successful anyway. Maybe Harvard itself had no effect on them. A few years ago, a psychotherapist named Mary Pipher wrote a bestseller called *Reviving Ophelia* (1994), in which she described the difficult lives of—as she saw it—typical adolescent girls. Pipher painted a stark picture of depression, rampant eating disorders, low self-esteem, academic failure, suicidal thoughts, and even suicide itself. Where did she get this picture? From her patients—that is, from adolescent girls who were in deep despair, or at least were unhappy enough to seek help. If Pipher had talked with a comparison sample of girls who hadn’t sought help, perhaps the story would not have been so bleak.

In the Sherman and Berk (1984) domestic violence experiment in Minneapolis, some police officers sometimes violated the random assignment plan when they thought the circumstances warranted arresting a suspect who had been randomly assigned to receive just a warning; thus, they created a selection bias in the experimental group.

- *Mortality*—Even when random assignment works as planned, the groups can become different over time because of **mortality, or differential attrition**; this can also be called “deselection.” That is, the groups become different because subjects are more likely to drop out of one of the groups for various reasons. At some colleges, satisfaction surveys show that seniors are more likely to rate their colleges positively than are freshmen. But remember that the freshmen who really hated the place may have transferred out, so *their* ratings aren’t included with senior ratings. In effect, the lowest scores are removed; that’s a mortality problem. This is not a likely problem in a laboratory experiment that occurs in one session, but some laboratory experiments occur over time, and so differential attrition can become a problem. Subjects who experience the experimental condition may become more motivated to continue in the experiment than comparison subjects.



Note that whenever subjects are not assigned randomly to treatment and comparison groups, the threat of selection bias or mortality is very great. Even if the comparison group matches the treatment group on important variables, there is no guarantee that the groups were similar initially in terms of either the dependent variable or some other characteristic. However, a pretest helps the researchers to determine and control for selection bias.

- *Instrument Decay*—Measurement instruments of all sorts wear out, producing different results for cases studied later in the research. An ordinary spring-operated bathroom scales, for instance, becomes “soggy” after some years, showing slightly heavier weights than would be correct. Or a college teacher—a kind of instrument for measuring student performance—gets tired after reading too many papers one weekend and starts giving everyone a B. Research interviewers can get tired or bored, too, leading perhaps to shorter or less thoughtful answers from subjects. In all these cases, the measurement instrument has “decayed,” or worn out.

Endogenous Change

The next three problems, subsumed under the label **endogenous change**, occur when natural developments in the subjects, independent of the experimental treatment itself, account for some or all of the observed change between pretest and posttest.

- *Testing*—Taking the pretest can itself influence posttest scores. As the Kaplan SAT prep courses attest, there is some benefit to just getting used to the test format. Having taken the test beforehand can be an advantage. Subjects may learn something or may be sensitized to an issue by the pretest and, as a result, respond differently the next time they are asked the same questions, on the posttest.
- *Maturation*—Changes in outcome scores during experiments that involve a lengthy treatment period may be due to maturation. Subjects may age, gain experience, or grow in knowledge—all as part of a natural maturational experience—and thus respond differently on the posttest than on the pretest. In many high school yearbooks, seniors are quoted as saying, for instance, “I started at West Geneva High as a boy and leave as a man. WGHS made me grow up.” Well, he probably would have grown up anyway, high school or not. WGHS wasn’t the cause.
- *Regression*—Subjects who are chosen for a study because they received very low scores on a test may show improvement in the posttest, on average, simply because some of the low scorers were having a bad day. Whenever



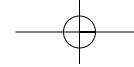
subjects are selected for study because of extreme scores (either very high or very low), the next time you take their scores they will likely “regress,” or move toward the average. For instance, suppose you give an IQ test to third graders and then pull the bottom 20% of the class out for special attention. The next time that group (the 20%) takes the test, they’ll almost certainly do better—and not just because of testing practice. In effect, they *can’t* do worse—they were at the bottom already. On average, they must do better. A football team that goes 0–12 one season almost has to improve. A first-time novelist writes a wonderful book, and gains worldwide acclaim and a host of prizes. The next book is not so good, and critics say “The praise went to her head.” But it didn’t; she *couldn’t* have done better. Whenever you pick people for being on an extreme end of a scale, odds are that next time they’ll be more average. This is called **regression**.

Regression effects: A source of causal validity that occurs when subjects who are chosen for a study because of their extreme scores on the dependent variable become less extreme on the posttest due to natural cyclical or episodic change in the variable.

Testing, maturation, and regression effects are generally not a problem in experiments that have a control group because they would affect the experimental group and the comparison group equally. However, these effects could explain any change over time in most before-and-after designs, because these designs do not have a comparison group. Repeated measures, panel studies, and time series designs are better in this regard because they allow the researcher to trace the pattern of change or stability in the dependent variable up to and after the treatment. Ongoing effects of maturation and regression can thus be identified and taken into account.

History

History, or **external events** during the experiment (things that happen outside the experiment), could change subjects’ outcome scores. Examples are newsworthy events that have to do with the focus of an experiment and major disasters to which subjects are exposed. If you were running a series of diversity workshops for some insurance company employees while the O. J. Simpson trial was taking place, for instance, participants’ thoughts on race relations at the end of the workshops may say less about you than about O. J. Simpson, or about their own relationship with the judicial system. This problem is often referred to as a **history effect**—history during the experiment, that is. It is a particular concern in before-and-after designs.



Causal conclusions can be invalid in some true experiments because of the influence of external events. For example, in an experiment in which subjects go to a special location for the treatment, something at that location unrelated to the treatment could influence these subjects. External events are a major concern in studies that compare the effects of programs in different cities or states (Hunt, 1985:276–277).

Contamination

Contamination occurs in an experiment when the comparison and treatment groups somehow affect each other. When comparison group members know they are being compared, they may increase their efforts just to be more competitive. This has been termed **compensatory rivalry**, or the **John Henry effect**, named after the “steel driving man” of the folk song, who raced against a steam drill in driving railroad spikes and killed himself in the process. Knowing that they are being denied some advantage, comparison group subjects may as a result increase their efforts to compensate. On the other hand, comparison group members may become demoralized if they feel that they have been left out of some valuable treatment and may perform worse than expected as a result. Both compensatory rivalry and demoralization thus distort the impact of the experimental treatment.

The danger of contamination can be minimized if the experiment is conducted in a laboratory, if members of the experimental group and the comparison group have no contact while the study is in progress, and if the treatment is relatively brief. Whenever these conditions are not met, the likelihood of contamination increases.

Treatment Misidentification

Sometimes the subjects experience a “treatment” that wasn’t intended by the researcher. The following are three possible sources of **treatment misidentification**:

Expectancies of experiment staff—Change among experimental subjects may be due to the positive expectancies of the experiment staff who are delivering the treatment rather than to the treatment itself. Even well-trained staff may convey their enthusiasm for an experimental program to the subjects in subtle ways. This is a special concern in evaluation research when program staff and researchers may be biased in favor of the program for which they work and are eager to believe that their work is helping clients. Such positive staff expectations thus create a **self-fulfilling prophecy**. However, in experiments on the effects of treatments such as medical drugs, **double-blind procedures** can be used: Staff delivering the treatments do not know which subjects are getting the treatment and which are receiving a placebo—something that looks like the treatment but has no effect.



Placebo effect—In medicine, a *placebo* is a chemically inert substance (a sugar pill, for instance) that looks like a drug but actually has no direct physical effect. Research shows that such a pill can actually produce positive health effects in two-thirds of patients suffering from relatively mild medical problems (Goleman, 1993:C3). In other words, if you wish that a pill will help, it often actually does.

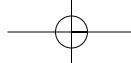
In social science research, such *placebo effects* occur when subjects think their behavior should improve through an experimental treatment and then it does—not from the treatment, but from their own belief. Researchers might then misidentify the treatment as having produced the effect.

Hawthorne effect—Members of the treatment group may change in terms of the dependent variable because their participation in the study makes them feel special. This problem could occur when treatment group members compare their situation to that of members of the control group who are not receiving the treatment, in which case it would be a type of contamination effect. But experimental group members could feel special simply because they are in the experiment. This is termed a **Hawthorne effect**, after a famous productivity experiment at the Hawthorne electric plant outside Chicago. No matter what conditions the researchers changed in order to improve or diminish productivity (for instance, increasing or decreasing the lighting in the plant), the workers seemed to work harder simply because they were part of a special experiment. Oddly enough, some more recent scholars suggest that in the original Hawthorne studies there was actually a selection bias, not a true Hawthorne effect—but the term has stuck (see Bramel & Friend, 1981). Hawthorne effects are also a concern in evaluation research, particularly when program clients know that the research findings may affect the chances for further program funding.

Process analysis is a technique for avoiding treatment misidentification (Hunt, 1985:272–274). Periodic measures are taken throughout an experiment to assess whether the treatment is being delivered as planned. For example, Drake et al. (1996) collected process data to monitor the implementation of two employment service models that they tested. One site did a poorer job of implementing the individual placement and support model than the other site, although the required differences between the experimental conditions were still achieved. Process analysis is often a special focus in evaluation research because of the possibility of improper implementation of the experimental program.

Generalizability

The need for generalizable findings can be thought of as the Achilles heel of true experimental design. The design components that are essential for a true



experiment and that minimize the threats to causal validity make it more difficult to achieve sample generalizability—being able to apply the findings to some clearly defined larger population—and cross-population generalizability—generalizing across subgroups and to other populations and settings.

Sample Generalizability

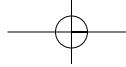
Subjects who can be recruited for a laboratory experiment, randomly assigned to a group, and kept under carefully controlled conditions for the duration of the study are unlikely to be a representative sample of any large population of interest to social scientists. Can they be expected to react to the experimental treatment in the same way as members of the larger population? The generalizability of the treatment and of the setting for the experiment also must be considered (Cook & Campbell, 1979:73–74). The more artificial the experimental arrangements, the greater the problem (Campbell & Stanley, 1966:20–21).

In some limited circumstances, a researcher may be able to sample subjects randomly for participation in an experiment and thus select a generalizable sample—one that is representative of the population from which it is selected. This approach is occasionally possible in **field experiments**. For example, some studies of the effects of income supports on the work behavior of poor persons have randomly sampled persons within particular states before randomly assigning them to experimental and comparison groups. Sherman and Berk's (1984) field experiment about the impact of arrest in actual domestic violence incidents (see Chapter 2) used a slightly different approach. In this study, all eligible cases were treated as subjects in the experiment during the data collection periods. As a result, we can place a good deal of confidence in the generalizability of the results to the population of domestic violence arrest cases in Minneapolis.

Cross-Population Generalizability

Researchers often are interested in determining whether treatment effects identified in an experiment hold true across different populations, times, or settings. When random selection is not feasible, the researchers may be able to increase the cross-population generalizability of their findings by selecting several different experimental sites that offer marked contrasts on key variables (Cook & Campbell, 1979:76–77).

Within a single experiment, researchers also may be concerned with whether the relationship between the treatment and the outcome variable holds true for certain subgroups. This demonstration of “external validity” is important evidence



about the conditions that are required for the independent variable(s) to have an effect. Price, Van Ryn, and Vinokur (1992) found that intensive job-search assistance reduced depression among individuals who were at high risk for it because of other psychosocial characteristics; however, the intervention did not influence the rate of depression among individuals at low risk for depression. This is an important limitation on the generalizability of the findings, even if the sample taken by Price et al. was representative of the population of unemployed persons.

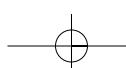
Finding that effects are consistent across subgroups does not establish that the relationship also holds true for these subgroups in the larger population, but it does provide supportive evidence. We have already seen examples of how the existence of treatment effects in particular subgroups of experimental subjects can help us predict the cross-population generalizability of the findings. For example, Sherman and Berk's research (see Chapter 2) found that arrest did not deter subsequent domestic violence for unemployed individuals; arrest also failed to deter subsequent violence in communities with high levels of unemployment.

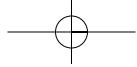
There is always an implicit tradeoff in experimental design between maximizing causal validity and generalizability. The more that assignment to treatments is randomized and all experimental conditions are controlled, the less likely it is that the research subjects and setting will be representative of the larger population. College students are easy to recruit and to assign to artificial but controlled manipulations, but both practical and ethical concerns preclude this approach with many groups and with respect to many treatments. However, although we need to be skeptical about the generalizability of the results of a single experimental test of a hypothesis, the body of findings accumulated from many experimental tests with different people in different settings can provide a very solid basis for generalization (Campbell & Russo, 1999:143).

Exhibit 5.7 Solomon Four-Group Design Testing the Interaction of Pretesting and Treatment

Experimental group:	R	O ₁	X	O ₂
Comparison group:	R	O ₁		O ₂
Experimental group:	R		X	O ₂
Comparison group:	R			O ₂

Key: R = Random assignment
 O = Observation (pretest or posttest)
 X = Experimental treatment





Interaction of Testing and Treatment

A variant on the problem of external validity occurs when the experimental treatment has an effect only when particular conditions created by the experiment occur. One such problem occurs when the treatment has an effect only if subjects have had the pretest. The pretest sensitizes the subjects to some issue so that when they are exposed to the treatment, they react in a way they would not have reacted if they had not taken the pretest. In other words, testing and treatment interact to produce the outcome. For example, answering questions in a pretest about racial prejudice may sensitize subjects so that when they are exposed to the experimental treatment, seeing a film about prejudice, their attitudes are different from what they would have been. In this situation, the treatment truly had an effect, but it would not have had an effect if it were repeated without the sensitizing pretest. This possibility can be evaluated by using the Solomon Four-Group Design to compare groups with and without a pretest (see Exhibit 5.7). If testing and treatment do interact, the difference in outcome scores between the experimental and comparison groups will be different for subjects who took the pretest compared to those who did not.

As you can see, there is no single procedure that establishes the external validity of experimental results. Ultimately, we must base our evaluation of external validity on the success of replications taking place at different times and places and using different forms of the treatment.

HOW DO EXPERIMENTERS PROTECT THEIR SUBJECTS?

Social science experiments often involve subject deception. Primarily because of this feature, some experiments have prompted contentious debates about research ethics. Experimental evaluations of social programs also pose ethical dilemmas because they require researchers to withhold possibly beneficial treatment from some of the subjects just on the basis of chance. Such research may also yield sensitive information about program compliance, personal habits, and even illegal activity—information that is protected from legal subpoenas only in some research concerning mental illness or criminal activity (Boruch, 1997). In this section, we will give special attention to the problems of deception and the distribution of benefits in experimental research.

Deception

Deception occurs when subjects are misled about research procedures in order to determine how they would react to the treatment if they were not research subjects. Deception is a critical component of many social experiments, in part because of the difficulty of simulating real-world stresses and dilemmas in



a laboratory setting. Stanley Milgram's (1965) classic study of obedience to authority provides a good example. Volunteers were recruited for what they were told was a study of the learning process. The experimenter told the volunteers they were to play the role of "teacher" and to administer an electric shock to a "student" in the next room when the student failed a memory test. The shocks were phony (and the students were actors), but the real subjects, the volunteers, didn't know this. They were told to increase the intensity of the shocks, even beyond what they were told was a lethal level. Many subjects continued to obey the authority in the study (the experimenter), even when their obedience involved administering what they thought were potentially lethal shocks to another person.

But did the experimental subjects actually believe that they were harming someone? Observational data suggest they did: "Persons were observed to sweat, tremble, stutter, bite their lips, and groan as they found themselves increasingly implicated in the experimental conflict" (Milgram 1965:66).

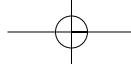
Verbatim transcripts of the sessions also indicated that participants were in much agony about administering the "shocks." So it seems that Milgram's deception "worked"; moreover, it seemed "necessary," since Milgram could not have administered real electric shocks to the students, nor would it have made sense for him to order the students to do something that wasn't so troubling, nor could he have explained what he was really interested in before conducting the experiment. The real question: Is this sufficient justification to allow the use of deception?

Aronson and Mills's study (1959) of severity of initiation (at an all-women's college in the 1950s) provides a very different example of the use of deception in experimental research—one that does not pose greater-than-everyday risks to subjects. The students who were randomly assigned to the "severe initiation" experimental condition had to read a list of embarrassing words. Even in the 1950s, reading a list of potentially embarrassing words in a laboratory setting and listening to a taped discussion were unlikely to increase the risks to which students were exposed in their everyday lives. Moreover, the researchers informed subjects that they would be expected to talk about sex and could decline to participate in the experiment if this requirement would bother them. No one dropped out.

To further ensure that no psychological harm was caused, Aronson and Mills (1959) explained the true nature of the experiment to the subjects after the experiment, in what is called **debriefing**. The subjects' reactions were typical:

None of the Ss expressed any resentment or annoyance at having been misled. In fact, the majority were intrigued by the experiment, and several returned at the end of the academic quarter to ascertain the result. (1959:179)

Although the American Sociological Association's *Code of Ethics* does not discuss experimentation explicitly, one of its principles highlights the ethical dilemma posed by deceptive research:



130 MAKING SENSE OF THE SOCIAL WORLD

(a) Sociologists do not use deceptive techniques (1) unless they have determined that their use will not be harmful to research participants; is justified by the study's prospective scientific, educational, or applied value; and that equally effective alternative procedures that do not use deception are not feasible, and (2) unless they have obtained the approval of institutional review boards or, in the absence of such boards, with another authoritative body with expertise on the ethics of research.

(b) Sociologists never deceive research participants about significant aspects of the research that would affect their willingness to participate, such as physical risks, discomfort, or unpleasant emotional experiences. (American Sociological Association, 1997:3)

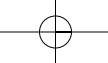
Selective Distribution of Benefits

Field experiments conducted to evaluate social programs also can involve issues of informed consent (Hunt, 1985:275–276). One ethical issue that is somewhat unique to field experiments is the **distribution of benefits**: How much are subjects harmed by the way treatments are distributed in the experiment? For example, Sherman and Berk's (1974) experiment, and its successors, required police to make arrests in domestic violence cases largely on the basis of a random process. When arrests were not made, did the subjects' abused spouses suffer? Price et al. (1992) randomly assigned unemployed individuals who had volunteered for job-search help to an intensive program. Were the unemployed volunteers who were assigned to the comparison group at a big disadvantage?

Is it ethical to give some potentially advantageous or disadvantageous treatment to people on a random basis? Random distribution of benefits is justified when the researchers do not know whether some treatment actually is beneficial or not—and, of course, it is the goal of the experiment to find out. Chance is as reasonable a basis for distributing the treatment as any other. Also, if insufficient resources are available to fully fund a benefit for every eligible person, distribution of the benefit on the basis of chance to equally needy persons is ethically defensible (Boruch, 1997:66–67).

CONCLUSION

Causation and the means for achieving causally valid conclusions in research is the last of the three legs on which the validity of research rests. In this chapter, you have learned about the five criteria used to evaluate the extent to which particular research designs may achieve causally valid findings. You have been exposed to the problem of spuriousness and the way that randomization deals with it. You also have learned why we must take into account the units of analysis in a research design in order to come to appropriate causal conclusions.

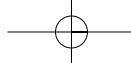


True experiments help greatly to achieve more valid causal conclusions—they are the “gold standard” for testing causal hypotheses. Even when conditions preclude use of a true experimental design, many research designs can be improved by adding some experimental components. However, although it may be possible to test a hypothesis with an experiment, it is not always desirable to do so. Laboratory experiments may be inadvisable when they do not test the real hypothesis of interest but test instead a limited version that is amenable to laboratory manipulation. It also does not make sense to test the impact of social programs that cannot actually be implemented because of financial or political problems (Rossi & Freeman, 1989:304–307). Yet the virtues of experimental designs mean that they should always be considered when explanatory research is planned.

We emphasize that understandings of causal relationships are always partial. Researchers must always wonder whether they have omitted some relevant variables from their controls or whether their experimental results would differ if the experiment were conducted in another setting or at another time in history. But the tentative nature of causal conclusions means that we must give more—not less—attention to evaluating the causal validity of social science research whenever we need to ask the simple question, “What caused variation in this social phenomenon?”

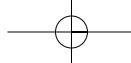
KEY TERMS

- | | |
|--|--|
| Association | Mechanism |
| Before-and-after design | Multiple group before-and-after design |
| Causal effect | Nonequivalent control group design |
| <i>Ceteris paribus</i> | Nonspuriousness |
| Compensatory rivalry (John Henry effect) | Placebo effect |
| Contamination | Posttest |
| Context | Pretest |
| Control group | Process analysis |
| Debriefing | Quasi-experimental design |
| Differential attrition (mortality) | Random assignment |
| Distribution of benefits | Randomization |
| Double-blind procedure | Regression effects |
| Endogenous change | Repeated cross-sectional design |
| Evaluation research | Repeated measures panel design |
| Ex post facto control group design | Selection bias |
| Expectancies of experimental staff | Self-fulfilling prophecy |
| Experimental group | Spurious relationship |
| External event | Time order |
| Field experiment | Time series design |
| Hawthorne effect | Treatment misidentification |
| History effect | True experiment |
| Matching | |



HIGHLIGHTS

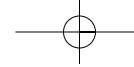
- Three criteria generally are viewed as necessary for identifying a causal relationship: association between the variables, proper time order, and nonspuriousness of the association. In addition, the basis for concluding that a causal relationship exists is strengthened by identification of a causal mechanism and the context.
- Association between two variables by itself is insufficient evidence of a causal relationship. This point is commonly made by the expression “Correlation does not prove causation.”
- The independent variable in an experiment is represented by a treatment or other intervention. Some subjects receive one type of treatment; others may receive a different treatment or no treatment. In true experiments, subjects are assigned randomly to comparison groups.
- Experimental research designs have three essential components: use of at least two groups of subjects for comparison, measurement of the change that occurs as a result of the experimental treatment, and use of random assignment. In addition, experiments may include identification of a causal mechanism and control over experimental conditions.
- Random assignment of subjects to experimental and comparison groups eliminates systematic bias in group assignment. The odds of there being a difference between the experimental and comparison groups on the basis of chance can be calculated. They become very small for experiments with at least 30 subjects per group.
- Random assignment and random sampling both rely on a chance selection procedure, but their purposes differ. Random assignment involves placing predesignated subjects into two or more groups on the basis of chance; random sampling involves selecting subjects out of a larger population on the basis of chance. Matching of cases in the experimental and comparison groups is a poor substitute for randomization because identifying in advance all important variables on which to make the match is not possible. However, matching can improve the comparability of groups when it is used to supplement randomization.
- Ethical and practical constraints often preclude the use of experimental designs.
- Quasi-experimental designs can be either a nonequivalent control group design or a before-and-after design. Nonequivalent control groups can be created through either individual matching of subjects or matching of group characteristics. In either case, these designs can allow us to establish the existence of an association and the time order of effects, but they do not ensure that some unidentified extraneous variable did not cause what we think of as the effect of the independent variable. Before-and-after designs can involve one or more pretests and posttests. Although multiple pretests and posttests make it unlikely that another, extraneous influence caused the experimental effect, they do not guarantee it.
- Ex post facto control group designs involve a comparison group that individuals could decide to join precisely because they prefer this experience rather than what the experimental group offers. This creates differences in subject characteristics between



the experimental and control groups that might very well result in a difference on the dependent variable. Because of this possibility, this type of design is not considered a quasi-experimental design.

- Invalid conclusions about causality may occur when relationships between variables measured at the group level are assumed to apply at the individual level (the ecological fallacy) and when relationships between variables measured at the level of individuals are assumed to apply at the group level (the reductionist fallacy). Nonetheless, many research questions point to relationships at multiple levels and may profitably be answered by studying different units of analysis.
- Causal conclusions derived from experiments can be invalid because of selection bias, endogenous change, the effects of external events, cross-group contamination, or treatment misidentification. In true experiments, randomization should eliminate selection bias and bias due to endogenous change. External events, cross-group contamination, and treatment misidentification can threaten the validity of causal conclusions in both true experiments and quasi-experiments.
- Process analysis can be used in experiments to identify how the treatment had (or didn't have) an effect—a matter of particular concern in field experiments. Treatment misidentification is less likely when process analysis is used.
- The generalizability of experimental results declines if the study conditions are artificial and the experimental subjects are unique. Field experiments are likely to produce more generalizable results than experiments conducted in the laboratory.
- The external validity of causal conclusions is determined by the extent to which they apply to different types of individuals and settings. When causal conclusions do not apply to all the subgroups in a study, they are not generalizable to corresponding subgroups in the population; consequently, they are not externally valid with respect to those subgroups. Causal conclusions can also be considered externally invalid when they occur only under the experimental conditions.
- Subject deception is common in laboratory experiments and poses unique ethical issues. Researchers must weigh the potential harm to subjects and debrief subjects who have been deceived. In field experiments, a common ethical problem is selective distribution of benefits. Random assignment may be the fairest way of allocating treatment when treatment openings are insufficient for all eligible individuals and when the efficacy of the treatment is unknown.

To assist you in completing the Web Exercises, please access the Study Site at <http://www.pineforge.com/mssw2> where you'll find the Web Exercises with accompanying links. You'll find other useful study materials like self-quizzes and e-flashcards for each chapter, along with a group of carefully selected articles from research journals that illustrate the major concepts and techniques presented in the book.



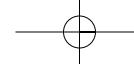
EXERCISES

Discussing Research

1. Review articles in several newspapers, copying down all causal assertions. These might range from assertions that the stock market declined because of uncertainty in the Middle East to explanations about why a murder was committed or why test scores are declining in U.S. schools. Inspect the articles carefully, noting all evidence used to support the causal assertions. Which of the five criteria for establishing causality are met? What other potentially important influences on the reported outcome have been overlooked? Can you spot any potentially spurious relationships?
2. Select several research articles in professional journals that assert, or imply, that they have identified a causal relationship between two or more variables. Are each of the criteria for establishing the existence of a causal relationship met? Find a study in which subjects were assigned randomly to experimental and comparison groups to reduce the risk of spurious influences on the supposedly causal relationship. How convinced are you by the study?
3. The practice CD-ROM contains lessons on units of analysis. Choose the Units of Analysis lesson from the main menu. It describes several research projects and asks you to identify the units of analysis in each.
4. The National Institutes of Health provides a tutorial for learning about current ethical standards in research. Complete this tutorial at <http://cme.nci.nih.gov/intro.htm>. Be prepared to spend one-half to one hour completing the tutorial. You must register as a college student and provide a bit of other information. Indicate that you do not need a certificate of completion. After you complete the registration fields, begin with the section on History. In this section, you will find a subsection on “The Development of Codes of Research Ethics.” When you get to the heading in this subsection on the “Belmont Report,” you will find a link to the federal “Common Rule” document. Click on this link and take the time to print the document out and read it. When you are finished with the tutorial and have read the Common Rule, you will be well on your way to becoming an expert on human subjects regulations. Identify the human subjects rules that are most important for research on human subjects.

Finding Research

1. Read an original article describing a social experiment. (Social psychology “readers,” collections of such articles for undergraduates, are a good place to find interesting studies.) Critique the article, using as your guide the article review questions presented in Exhibit 10.2. Focus on the extent to which experimental conditions were controlled and the causal mechanism was identified. Did inadequate control over conditions or inadequate identification of the causal mechanism make you feel uncertain about the causal conclusions?
2. Go to the Web site of the Community Policing Consortium at www.communitypolicing.org/about2.html. What causal assertions are made? Pick one of these assertions and propose a research design with which to test this assertion. Be specific.
3. Go to Sociosite at www.pscw.uva.nl/sociosite. Choose “Subject Areas.” Choose a sociological subject area you are interested in. Find an example of research that has been



done using experimental methods in this subject. Explain the experiment. Choose at least five of the Key Terms listed at the end of this chapter that are relevant to and incorporated in the research experiment you have located on the Internet. Explain how each of the five Key Terms you have chosen plays a role in the research example you found on the Web.

Critiquing Research

1. From newspapers or magazines, find two recent studies of education (reading, testing, etc.). For each study, list in order what you see as the most likely sources of internal invalidity (e.g., selection, mortality, etc.).
2. Select a true experiment, perhaps from the *Journal of Experimental and Social Psychology*, the *Journal of Personality and Social Psychology*, or sources suggested in class. Diagram the experiment using the exhibits in this chapter as a model. Discuss the extent to which experimental conditions were controlled and the causal mechanism was identified. How confident can you be in the causal conclusions from the study, based on review of the threats to internal validity discussed in this chapter: selection bias, endogenous change, external events, contamination, and treatment misidentification? How generalizable do you think the study's results are to the population from which the cases were selected? To specific subgroups in the study? How thoroughly do the researchers discuss these issues?
3. Repeat the previous exercise with a quasi-experiment.
4. Critique the ethics of one of the experiments presented in this chapter, or some other experiment you have read about. What specific rules do you think should guide researchers' decisions about subject deception and the selective distribution of benefits?

Doing Research

1. Try out the process of randomization. Go to the Web site www.randomizer.org. Now just type numbers into the randomizer for an experiment with two groups and 20 individuals per group. Repeat the process for an experiment with four groups and 10 individuals per group. Plot the numbers corresponding to each individual in each group. Does the distribution of numbers within each group truly seem to be random?
2. Participate in a social psychology experiment on the Internet. Go to www.socialpsychology.org/expts.htm. Pick an experiment in which to participate and follow the instructions. After you finish, write a description of the experiment and evaluate it using the criteria discussed in the chapter.
3. Volunteer for an experiment. Contact the psychology department and ask about opportunities for participating in laboratory experiments. Discuss the experience with your classmates.

Capítulo 4

Teoría de Causalidad

Capítulo 4

Teoría de Causalidad

Los objetivos de este capítulo son:

- Conocer todos los conceptos relacionados a la Teoría de Causalidad.
- Aplicar estos conceptos en ejemplos simples.

4.1 Introducción

La teoría de la causalidad constituye una de las herramientas básicas al momento de analizar la dinámica de los sistemas. La misma se basa en los siguientes tres aspectos:

- la causalidad,
- la retroalimentación y
- la frontera.

La causalidad tiene que ver con la influencia que ejerce un elemento del sistema sobre otro. La retroalimentación que trata de los sistemas dinámicos, los sistemas de ciclo cerrado. En la retroalimentación encontramos dos tipos: La positiva aquella en la cual una o más variables se alimentan continuamente a sí mismas para contribuir a su crecimiento o colapso y la retroalimentación negativa la que se caracteriza por un comportamiento dirigido u orientado hacia una meta. Y la frontera en la que se identifica el problema en estudio, se establecen las variables principales de estudio y esto dependiendo de a quienes les interesa el problema bajo consideración. Esta es una de las cosas más difíciles de hacer debido a la interrelación de un sistema con el otro. Estos diagramas presentan algunos inconvenientes: sólo son útiles durante las fases de conceptualización del modelo debido a que carecen de precisión y detalle de los datos estadísticos, niveles y cantidad de elementos auxiliares que se encuentran en el diagrama. Además, dichos diagramas ocultan información necesaria para la comprensión del comportamiento para una estructura de ciclo de retroalimentación.

A pesar de las desventajas antes descritas los diagramas de ciclos causales son muy utilizados, debido a que simplifican el estudio de las variables que intervienen en un determinado sistema y ayudan a conocer cuáles son las acciones que afectan a estas variables y cuáles son las causas o reacciones que las mismas proporcionan.

4.2 Causalidad o Causa-Efecto

El pensamiento causal es la clave para organizar las ideas en el campo de la dinámica de sistemas. Sin embargo, el concepto de causalidad puede ser muy sutil y la utilización del mismo requiere atención muy cuidadosa. Su notación es parecida a la que se presenta a continuación¹:



Dentro de un ciclo causal, una causa eventualmente se convierte en un efecto y así sucesivamente. Este proceso se le conoce como el proceso de retroalimentación². Para ilustrar mejor este punto presentaremos a continuación un ejemplo de lo que sería el sistema muy simplificado del ciclo del agua en nuestro planeta.

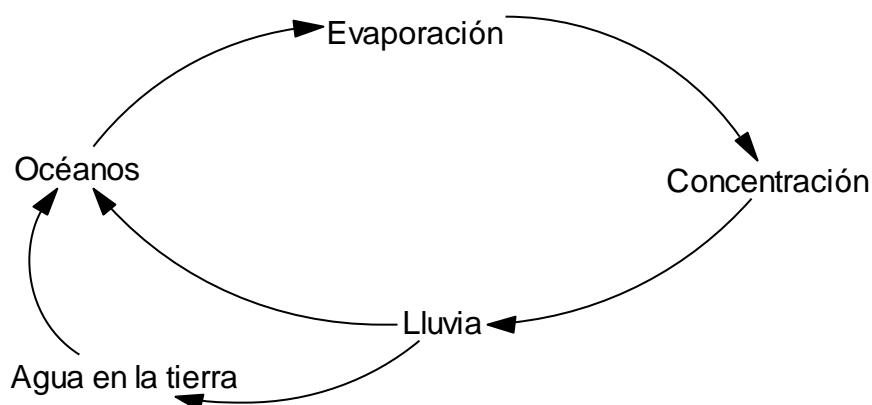
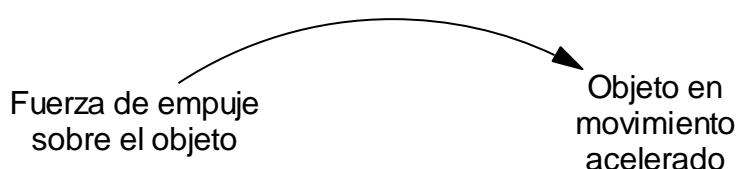


Figura 4.1: Diagrama sencillo del Ciclo del Agua en el Planeta Tierra

Los mejores ejemplos de causalidad involucran leyes físicas. Una fuerza aplicada sobre un objeto en descanso mueve el objeto en la dirección de la fuerza y una fuerza continua aplicada al objeto causa una aceleración continua:

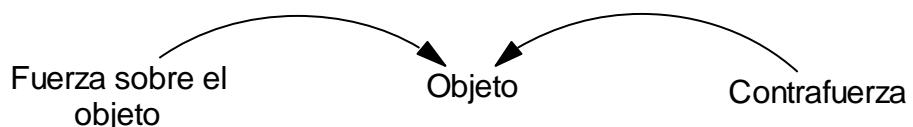


Un individuo ejerciendo una fuerza sobre un edificio, lo más seguro es que no lo moverá. Cuando el individuo ejerce la fuerza sobre el edificio, todas las otras cosas no se

¹ La flecha tiene que ser curva, ya que la flecha recta significa *infiere*.

² retroalimentación

mantuvieron igual. La madera, el acero, el concreto o combinaciones de estas que componen la pared del edificio ejercen contra fuerzas que son mayores a la del individuo. Para diagnosticar correctamente una influencia causal se debe ejecutar el experimento mental de preguntarse qué pasaría si la influencia causal bajo consideración fuera la única influencia que actúa sobre el objeto afectado.



Otro ejemplo: Los nacimientos infantiles hacen que la población crezca.

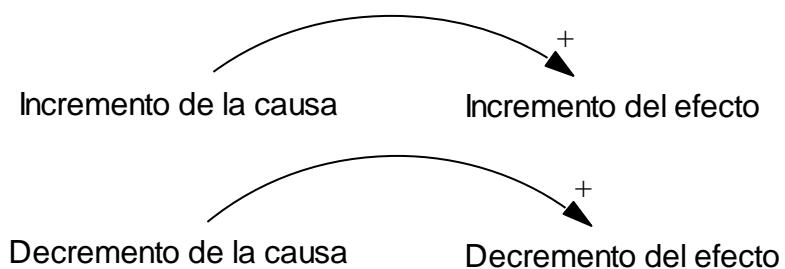


Esta causa y efecto es correcta, aunque países con tasa de nacimiento positivo tienen un crecimiento negativo de la población o sea que los nacimientos no son los únicos que afectan la población. Otras causas que afectan la población son: Muertes por nacimiento, accidentes, enfermedades, suicidios, guerras, migración, etc.

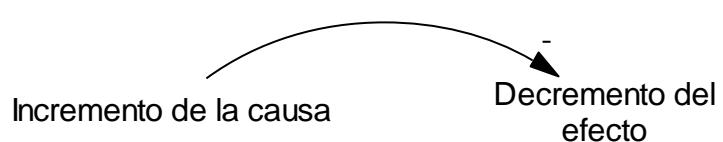
4.3 Polaridad de las Causas Efectos

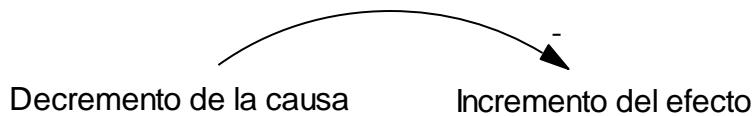
Toda causa-efecto tiene una polaridad positiva o negativa. Esta polaridad se coloca en la cabeza de la flecha.

Es positiva si se presenta un:



Será negativa si se presenta un:

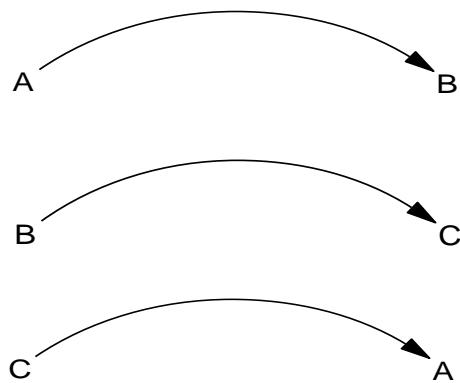




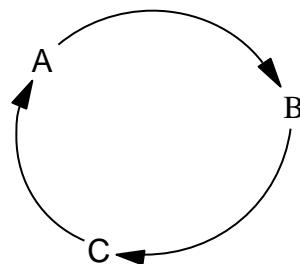
4.4 Ciclos Causales

4.4.1 Definición

Todo problema bajo investigación arroja un conjunto de causas-efectos que pueden unirse y formar ciclos de causas-efectos conocidos como ciclos causales. Por ejemplo, tenemos tres causas-efectos que el modelador descubre para un problema y son:



Estas tres causas-efectos se pueden unir y forman un ciclo causal:

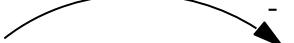
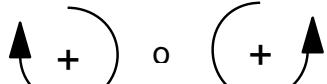


Cada problema tendrá el número de ciclos causales que se determinen y a esto lo llamamos diagrama de ciclos causales, una forma conceptual de representar o abstraer la realidad del sistema y presentarla en un modelo que es una abstracción del sistema dinámico bajo estudio. Los sistemas dinámicos enfocan la estructura y el comportamiento de los sistemas a través de la interacción de los ciclos de retroalimentación. Los Diagramas de Ciclos Causales identifican los principales ciclos de retroalimentación sin distinguir entre la naturaleza de las variables interconectadas.

4.4.2 Símbología

A continuación, se presenta un cuadro con todos los componentes de una causa-efecto. La primera columna presenta el símbolo y la segunda presenta el significado.

Tabla Nº 1

Símbolo	Significado
	La flecha curva ⁱ es usada para indicar la causalidad. El elemento en la cola de la flecha causa un cambio en el elemento a la cabeza de la flecha.
	Variables que deben estar muy bien definidas para evitar ambigüedades. Cada variable debe tener un nombre largo, un nombre corto o abreviado y una definición completa y clara. Es importante definir variables cuantitativas y cualitativas.
	El signo positivo cerca de la cabeza de la flecha indica que el elemento en la cola de la flecha y el elemento a la cabeza de la flecha cambian en la misma dirección. Si el elemento en la cola incrementa, el elemento en la cabeza incrementa; si el elemento en la cola decrementa, el elemento en la cabeza decrementa.
	El signo negativo cerca de la cabeza de flecha indica que el elemento en la cola de la flecha y el elemento a la cabeza de la flecha cambian en direcciones opuestas. Si el elemento en la cola incrementa, el elemento en la cabeza decrementa; si el elemento en la cola decrementa, el elemento en la cabeza incrementa.
	Este símbolo, encontrado en el medio de un ciclo cerrado, indica que el ciclo continúa en la misma dirección, causando un crecimiento sistemático o un colapso (caída), este movimiento la aleja del punto de equilibrio. Esto es llamado un ciclo de retroalimentación positiva.
	Este símbolo, encontrado en el medio de un ciclo cerrado, indica que el ciclo cambia de dirección, causando que el sistema fluctúe o se mueva hacia el equilibrio. Esto es llamado un ciclo de retroalimentación negativo.

Algunas veces las flechas (hipótesis o causa-efectos) tienen más sentido si se usa la palabra afecta o influye en vez de causa. Por ejemplo, el consumo de alimentos influye

en el peso, tiene más sentido que el consumo de alimentos causa el peso. Las palabras *causa*, *afecta* e *influye*, que se usan aquí, tienen aproximadamente el mismo significado.

Por ejemplo, la figura 4.2 presenta la siguiente causa efecto: Si el consumo de alimento aumenta, entonces el peso de la persona aumenta. Es lo mismo que: El consumo de alimentos afecta el peso de la persona.



Figura 4.2: Ejemplo de Causa Efecto

4.4.3 Usos y Beneficios de las Causas-Efectos

Como veremos en el próximo capítulo, una serie de causas efectos sobre un problema específico se une y forma lo que conoceremos como un diagrama de ciclos causales, donde las causas-efectos forman ciclos. Los diagramas de ciclos causales juegan dos papeles importantes en el estudio de los sistemas dinámicos. Primero, durante el desarrollo del modelo, sirven de bosquejos preliminares de las hipótesis causales. Segundo, los diagramas de ciclos causales pueden simplificar la ilustración de un modelo. En ambos sentidos, los diagramas de ciclos causales permiten al analista comunicar rápidamente la proposición estructural fundamental de su modelo.

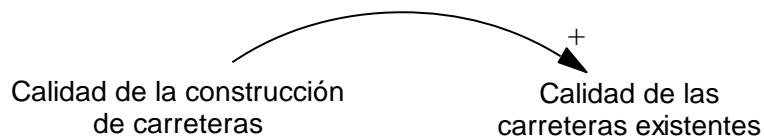
La diagramación de los ciclos causales estimula al modelador a conceptualizar los modelos reales del mundo en términos de ciclos de retroalimentación. Estos diagramas constituyen un instrumento de comunicación conveniente, sin embargo, los mismos presentan inconvenientes como veremos a continuación: una vez el diagrama de ciclo causal ha sido debidamente validado, entonces podemos convertirlo a un diagrama de flujos de datos (DFD) que veremos en el capítulo 6 el cual puede codificarse en lenguajes de simulación como Stella y Vensim.

4.5 Otros Ejemplos

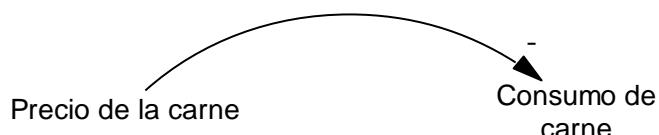
Si el número de kilómetros de carretera construidos por año aumenta, entonces también aumenta el número de kilómetros de carretera existente aumentará:



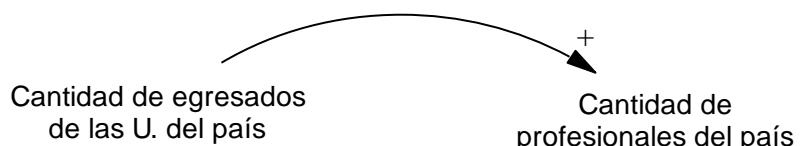
Si la calidad de construcción de las carreteras aumenta, aumenta la calidad de todas las carreteras existentes:



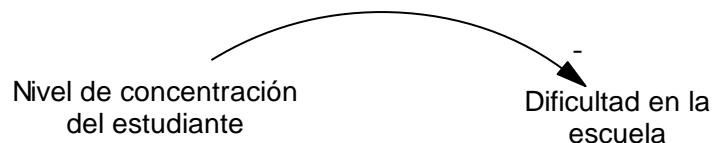
Si el precio de la carne aumenta, disminuye la cantidad de libras de carne consumidas:



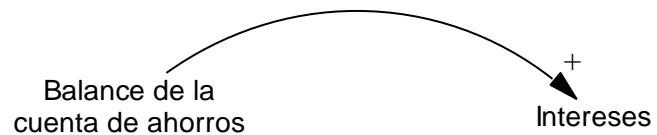
Si la cantidad de egresados de las universidades del país aumenta, aumenta la cantidad de profesionales en todo el país:



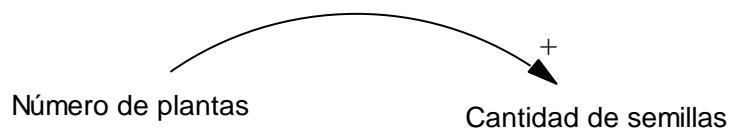
Si el nivel de concentración del estudiante incrementa, también decrementará la dificultad que él confronta:



Si el balance en la cuenta de ahorros aumenta, aumentan los intereses pagados en la cuenta de ahorros:



Si el número de plantas aumenta, aumenta la cantidad de semillas que ellas producen:



4.6 Retroalimentación³

A continuación, se presenta el concepto de *retroalimentación*, que es el concepto más importante de la teoría causal. Existen dos tipos de retroalimentación: la *positiva* y la *negativa*.

4.6.1 Positiva

En este proceso una variable se alimenta continuamente a sí misma para contribuir a su crecimiento o colapso. Los elementos básicos de un sistema de retroalimentación positiva son:

- Una acción
- Un refuerzo de la acción

A continuación, una gráfica que detalla, en sí, tales elementos:

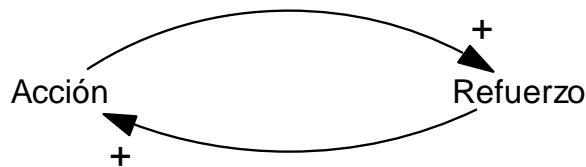


Figura 4.6 Retroalimentación Positiva

4.6.1.1 Retroalimentación Positiva Viciosa

En un círculo vicioso, el empeoramiento de un elemento en la cadena causal ocasiona una mayor degradación del elemento. Un ejemplo clásico del círculo vicioso es el de la espiral inflacionaria del precio-salario. Los trabajadores, inquietos por los altos precios, demandan salarios superiores. Los fabricantes aumentan sus precios para cubrir el costo de la mano de obra. Los trabajadores perciben que los precios suben otra vez y demandan salarios superiores para cubrir estos incrementos. Esta demanda adicional cierra el ciclo. Dentro del ejemplo anterior, un incremento en una variable conlleva al incremento de las otras tres. Este incremento posterior lleva al incremento de la primera variable y continua así. Las cuatro variables se realizan mutuamente y tienden a un

³ O *realimentación*.

incremento sin límite. El ciclo positivo en el centro del ciclo simboliza la naturaleza de retroalimentación positiva del ciclo.

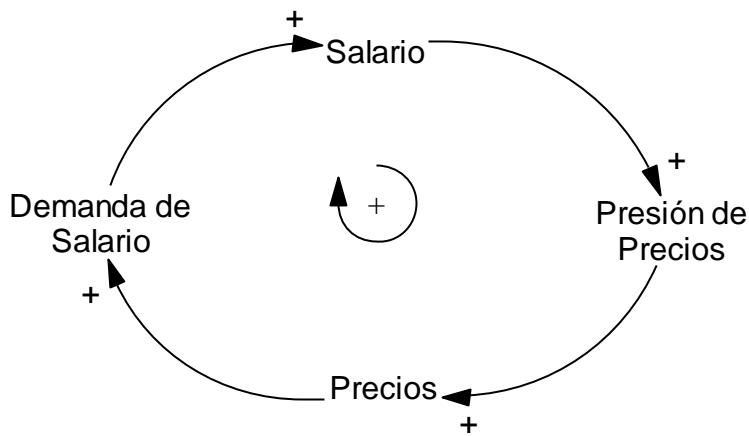


Figura 4.3: Retroalimentación Positiva Precio-Salario

4.6.1.2 Retroalimentación Positiva Virtuosa

En un ciclo virtuoso, un cambio positivo en un elemento del sistema adiciona un mejoramiento en él. Un ejemplo de este tipo de retroalimentación se presenta cuando existe un incremento de las exportaciones efectuadas en un país, las cuales dan como resultado un aumento de las ventas y al mismo tiempo un mejoramiento en la economía del país.

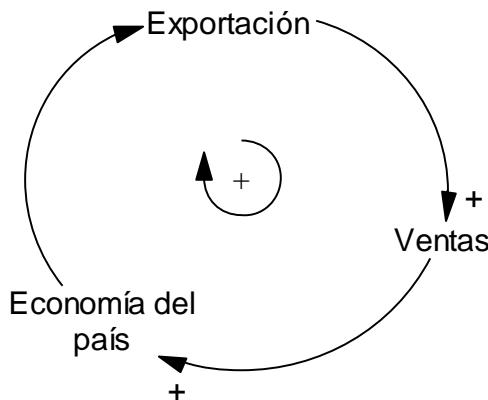


Figura 4.4: Retroalimentación Positiva Exportación-Economía del País

4.6.1.3 Modelo de Colapso Exponencial

Es una forma no muy común de retroalimentación positiva. La población del mundo, la producción de alimentos, la industrialización, la contaminación y el consumo de recursos naturales no renovables, exhiben un crecimiento exponencial. Tal crecimiento, que parece mínimo al principio, sube precipitadamente en un corto período de tiempo.

El ejemplo del asistente del joyero provee otra reveladora ilustración de la naturaleza explosiva del crecimiento exponencial. Un asistente ofreció trabajar para el joyero por un año con la condición que su salario empezara desde 0.2 centavos y se doblase cada semana. El joyero, contento con el pensamiento de una mano de obra barata, enseguida accedió. A la quinta semana, la paga del asistente era de 3.2 centavos; la décima 102.4 Balboas; y la decimoquinta, era 3,355.44 Balboas. Dos semanas después, el joyero tuvo que vender su tienda sólo para pagar a su asistente.

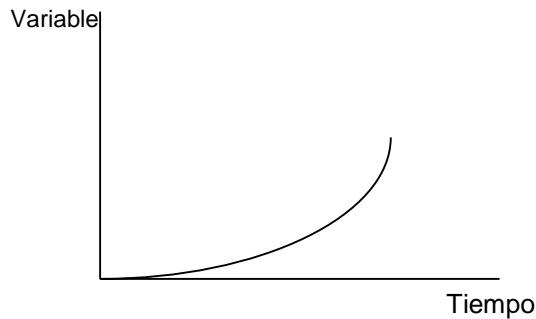


Figura 4.5: Curva de Crecimiento Exponencial

4.6.2 Retroalimentación Negativa

La retroalimentación negativa está caracterizada por un comportamiento dirigido u orientado hacia una meta. Términos tales como auto-gobernación, autorregulación, auto equilibrio, servomecanismo o adaptable, todos implican la presencia de una meta y definen sistemas de retroalimentación negativa.

4.6.2.1 Ejemplo

El calentador termostáticamente controlado es un sistema común de auto - gobernación. Un diagrama de ciclo causal del sistema es mostrado en la figura 4.6. El sistema del termostato procura mantener la temperatura del cuarto a una que ha sido preseleccionada. La unidad que toma las decisiones el termostato, detecta una diferencia entre la temperatura deseada y la temperatura actual de la habitación y activa la unidad de calentamiento. El calor adicional aumenta eventualmente la temperatura del cuarto al nivel deseado. Luego el termostato automáticamente desactiva la calefacción.

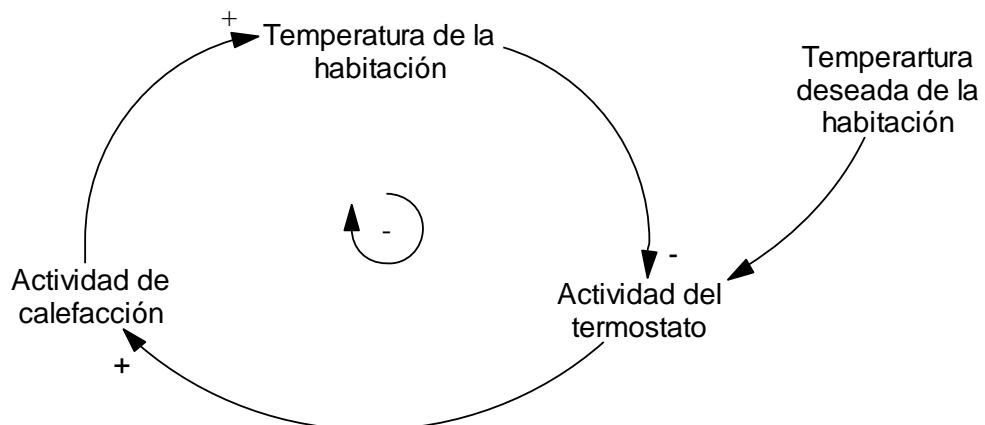


Figura 4.6: Sistema de Calefacción

4.6.2.2 Objetivos de la Retroalimentación

Estos sistemas desarrollados en el campo de la ingeniería de control, persiguen y tienden a mantener objetivos específicos. El concepto de control, en sí mismo, involucra la orientación hacia una meta. La acción dirigida a un objetivo es fundamental en el comportamiento social humano. Considere la socialización de un niño en casa. Los padres transmiten sus valores, actitudes y esperanzas al niño a través de un proceso de retroalimentación negativa. Cuando surge una discrepancia entre el comportamiento deseado y el comportamiento percibido del niño, los padres inician la acción correctiva en forma de castigo. El niño, a su vez, aprende como aplacar a sus padres y/o obtener la recompensa deseada. "El comportamiento con los cuales los padres controlan a sus hijos, ya sea en forma adversa o a través de refuerzos positivos, es moldeado y mantenido por la reacción del niño".

4.6.2.3 Elementos Básicos

Los cuatro elementos básicos de un sistema de retroalimentación negativa son:

- El estado deseado o meta.
- La discrepancia.
- La acción.
- El estado del sistema o nivel.

En la figura 4.7 se presenta un diagrama con estos 4 elementos básicos:

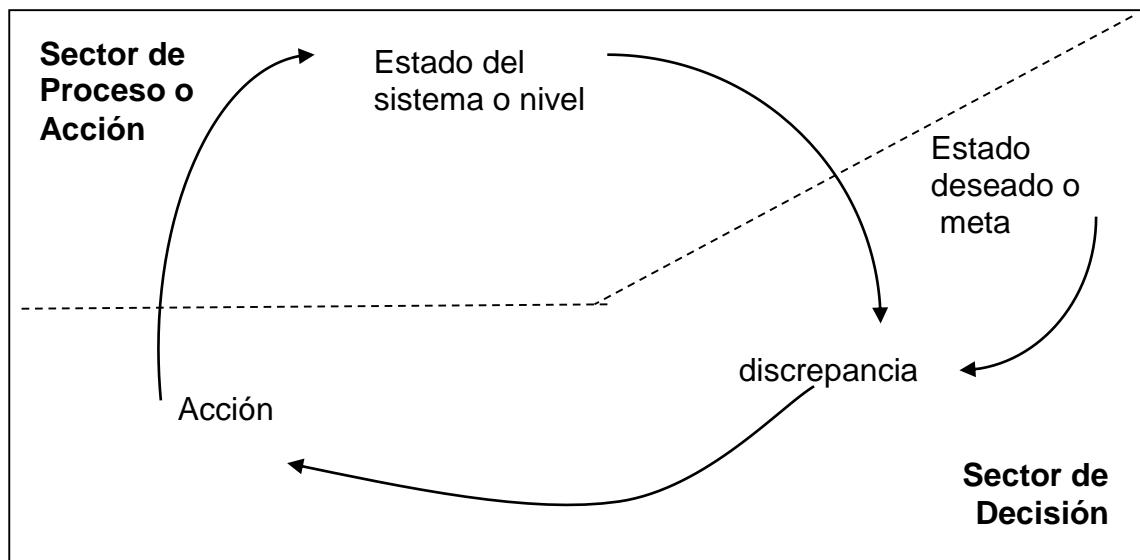


Figura 4.7: Elementos Básicos de un Ciclo de Retroalimentación Negativo⁴

La diferencia entre el ciclo de retroalimentación negativa y el ciclo de retroalimentación positiva, reside completamente en el sector de proceso de acción: la información percibida, la comparación, y los componentes de la toma de decisiones intervienen entre el estado del sistema y la acción. Sistemas sin un sector de acción son sistemas de "ciclo abierto". La acción tomada no depende del estado. El sector del proceso de decisión especifica cómo el estado del sistema controla la acción "... el proceso de decisión es aquel que controla cualquier acción del sistema. Puede ser una acción subconsciente, pueden ser los procesos gobernantes en el desarrollo biológico, pueden ser las consecuencias naturales de la estructura física del sistema". El más simple sistema de retroalimentación negativa contiene una meta o un elemento de discrepancia en el sector de proceso de decisión. La meta sirve como una referencia o línea de guía en las cuales se basa la acción del sistema. La meta es determinada externamente o exógenamente. Es decir, un enlace causal directo entre el sistema y la meta no se encuentra. La discrepancia entre la meta y el estado del sistema determina la magnitud y la dirección de la acción correctiva tomada. Por ejemplo, el termostato ejecuta esta función en el sistema de calefacción. Un censor, supervisa la temperatura del cuarto y la compara con la preseleccionada en el termostato. Una discrepancia de temperatura causa que el termostato active o desactive la unidad de calefacción.

El sector de proceso de decisión en el hombre es el cerebro. Por ejemplo, a través de la percepción visual, un conductor siente la posición de su vehículo relativo a la acera. La acción correctiva ocurre cuando el automóvil se mueve a una posición no deseada. La alternativa mecánica del volante, modifica la posición del automóvil. La figura 4.7 ilustra como el sector de decisión de la retroalimentación negativa completa el círculo del

⁴ Forrester Jay, "Principles of Systems".

sistema y actúa junto con los procesos de acción para producir el comportamiento buscado de la meta. Un incremento o decremento arbitrario en el estado del sistema produce una discrepancia entre la meta y el nivel del sistema. Para minimizar la discrepancia, el sistema inicia la acción para decrementar (incrementar) el nivel.

4.7 Ecuaciones de las Causa-Efectos

La causa-efecto:



Se traduce a:

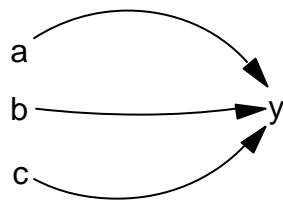
$y = f(x)$ y es la causa-efecto más sencilla que puede ocurrir.

Ahora bien, matemáticamente cómo es la fórmula, depende de la regresión lineal o no lineal que existe entre la variable independiente y la variable dependiente. Para determinar la ecuación, tenemos varias alternativas. La más común, consiste en recolectar datos de las dos variables y utilizar una herramienta estadística como SAS o SPSS que determina la ecuación. El resultado puede ser:

$$y = 2.0341 + 21.095x \quad o$$

$$y = 2.056x^2$$

Otro ejemplo de causa-efecto:



Este ejemplo se traduce:

$$Y = f(a,b,c)$$

Para determinar esta ecuación, se requiere de regresión múltiple y de los datos de las tres variables independientes y de la variable dependiente. Un paquete estadístico, como SAS, examina los datos de las cuatro variables y determina la ecuación que más se ajusta a los datos (para eso, SAS examina muchas alternativas). La ecuación puede ser polinomial, exponencial, trigonométrica, entre otras. El resultado puede ser:

$$y = 21.3 + 1.2a - 0.35b^2 + 4.07c^3 \quad o$$

$$y = -8.7 + 2.6\cos(a) - 6.1\log(b) + 4.3c^e$$

Cada causa-efecto(s) posee una ecuación única que explica el comportamiento entre las variables independientes y la variable dependiente. El modelador tiene que determinar la ecuación con la mayor exactitud posible y este puede consumir mucho esfuerzo y dinero.

4.8 Frontera

El pensamiento orientado a la causalidad ayuda a establecer las relaciones que se presentan entre las variables del sistema. Sin embargo, también hay que poner mucha atención al establecimiento de la frontera del mismo. Principalmente se debe identificar apropiadamente el problema en estudio, estableciendo cuales son las variables principales del estudio y esto de acuerdo a quienes les interesa este problema bajo consideración. Esta es una de las cosas más difíciles de hacer debido a la interrelación de un sistema con el otro.

A continuación, se presentan algunas observaciones para el reconocimiento de la frontera del sistema:

- Las variables tienen que estar altamente relacionadas y son necesarias para entender el sistema.
- En el momento que una variable importante del sistema se relaciona con otra que lleva al sistema a otro tema, entonces deténgase.

4.9 Diagramas de Ishikawa

Este diagrama es una variación interesante donde se tienen muchas causas y un solo efecto importante bajo estudio. Esta técnica es una forma de analizar problemas complejos que aparentan tener muchas causas interrelacionadas que afectan a un solo problema. Uno de los aspectos clave de esta técnica es el uso de diagramas causa – efecto. Debido a que la apariencia del diagrama, la técnica también es conocida como diagrama de espina de pescado⁵. Los beneficios de esta técnica incluyen:

- Permite explorar varias categorías de causas para un solo efecto.
- Fomenta la creatividad a través del proceso de lluvia de ideas.
- Proporciona una imagen visual del problema y las categorías de las causas potenciales.

⁵ Otro nombre que se puede escuchar para esta técnica es el de Diagramas de Ishikawa, debido a que el Dr. Japonés Kaoru Ishikawa fue el primero en usarla en 1943.

Los cinco pasos de Ishikawa:

- **Identificar el Problema**

Lo primero que se hace consiste en definir correctamente el problema que es el efecto final bajo estudio.

- **Presentar Problema y la Espina Dorsal**

Se dibuja una larga flecha horizontal apuntando hacia la caja. Esta flecha servirá como espina dorsal en la cual causas mayores y menores serán categorizadas e interrelacionadas.

- **Especificar las Grandes Categorías (Causas Mayores)**

Identificar causas potenciales y agruparlas en categorías mayores. Ejemplos de categorías mayores incluyen Gente, Procesos, Material, Equipo, Ambiente, etc. Las categorías de mayor nivel son identificadas a través de técnicas de lluvia de ideas, así que en este punto la preocupación no es si se está de acuerdo respecto a si una categoría contiene o no la causa potencial. Solo hay que ponerlas todas. Se debe asegurar que se deja el espacio suficiente entre las categorías de mayor nivel de tal forma que se pueda agregar información de menor nivel de detalle posteriormente.

- **Identificar las Causas Menores**

Se continúa con la lluvia de ideas de las causas al analizar más detalladamente las explicaciones de cada una de las categorías de mayor nivel, identificadas anteriormente. Se deben escribir causas más detalladas en líneas perpendiculares que se unan a las barras de la causa de mayor nivel, según sea el caso. Algunas veces, las causas detalladas podrán tener un mayor nivel de detalle, si eso sucede, será necesario conectar líneas adicionales a las causas correspondientes según sea el caso.

- **Identificar Acciones Correctivas**

Finalmente, se genera un plan de acción para resolver el problema. Hay que recordar que puede haber un gran número de causas potenciales que interactúan en conjunto para crear el problema. El plan de acción deberá tomar en cuenta estas interdependencias.

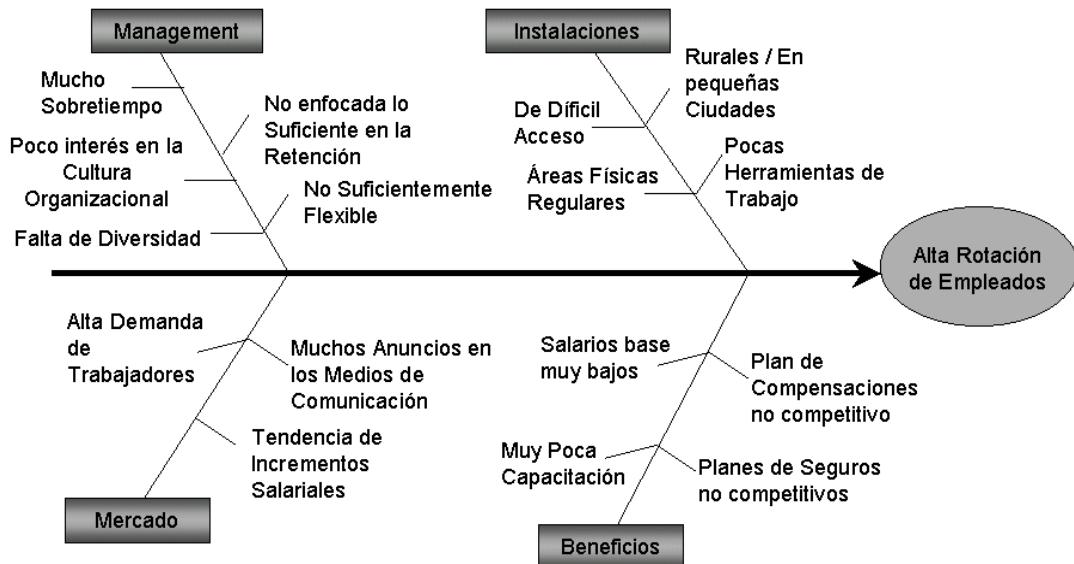


Figura 4.8: Ejemplo de un Diagrama de Ishikawa

Es muy oportuno en este capítulo incluir el siguiente diagrama simplificado de la Calidad de la Educación desarrollado por el autor (el diagrama original contiene más variables):

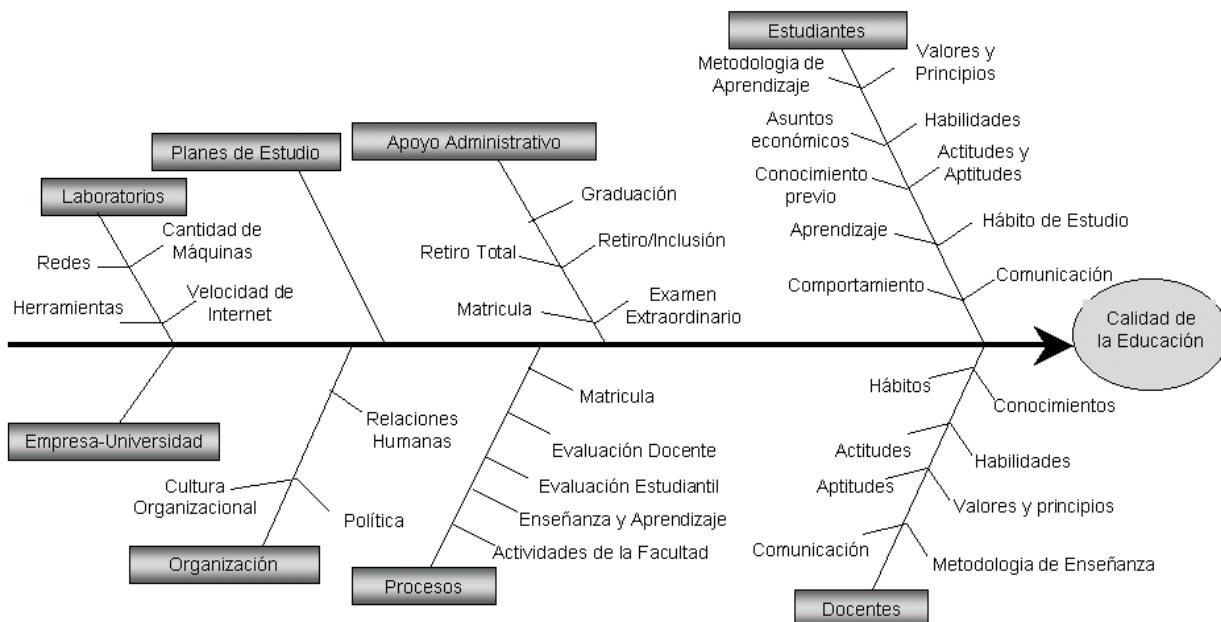


Figura 4.9: Ejemplo Simplificado de la Calidad de la Educación

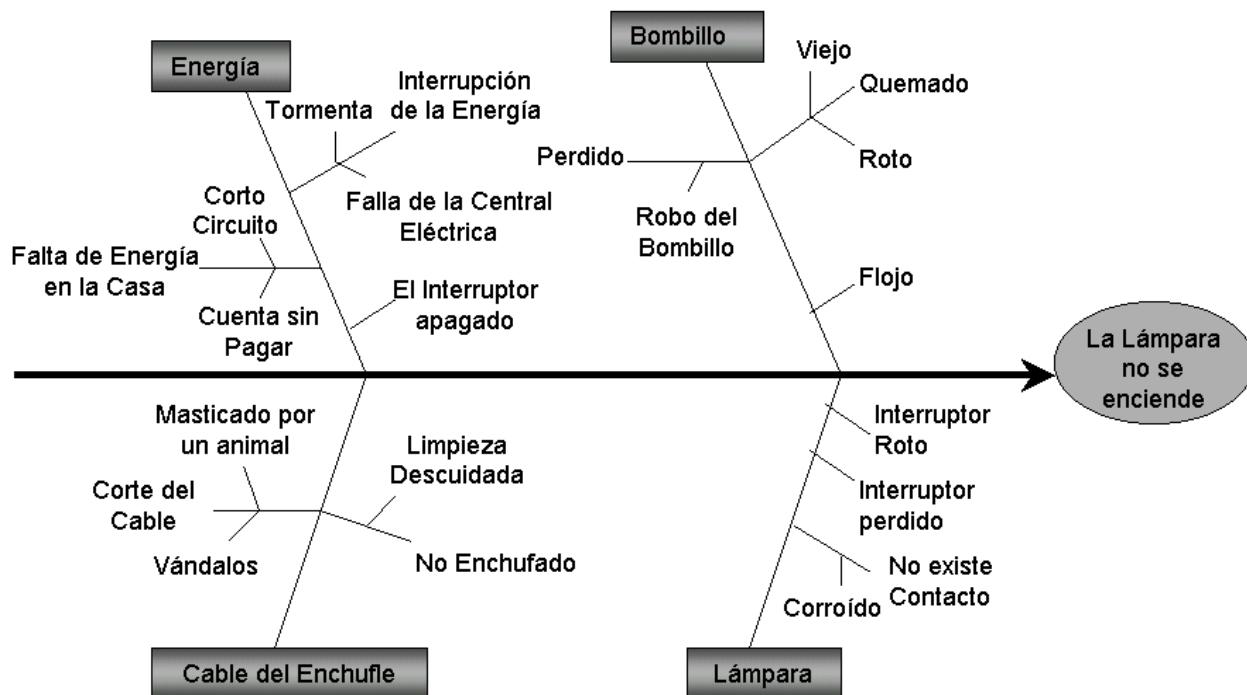
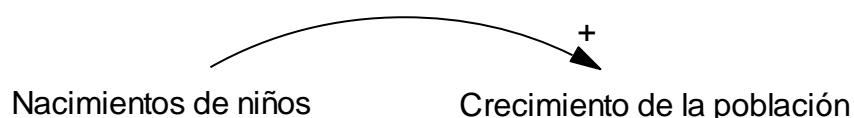


Figura 4.10: Ejemplo de la Lámpara que no enciende

4.10 Resumen

Los ciclos causales son importantes en el estudio del comportamiento de los sistemas en lo que a sistemas dinámicos se refiere. Estos proporcionan un bosquejo preliminar a las hipótesis causales, además de que simplifican la ilustración de un modelo. Los ciclos de causalidad, como también se les conocen, permiten al analista comunicar rápidamente la dimensión y estructura de su modelo de sistema. La asignación de las polaridades de los ciclos causales, las cuales se entienden como un tipo de nomenclatura que se utiliza en la diagramación de estos ciclos, las cuales nos indican como es el comportamiento de una determinada causa-efecto dependiendo de ciertos patrones de crecimiento o decremento. Por ejemplo, si hablásemos de la relación que existe entre las variables: Nacimiento de niños y crecimiento de la población, podemos decir que mientras mayor sea la cantidad de nacimientos el crecimiento de la población aumentaría, esto sin considerar que existen otras variables que podrían afectar el crecimiento de la población de una manera positiva o negativa. La representación de la relación antes descrita quedaría como sigue a continuación:



Dependiendo de la polaridad los ciclos de retroalimentación pueden ser clasificados como ciclos de retroalimentación positiva y ciclos de retroalimentación negativa. Se entiende por ciclos de retroalimentación positiva aquellos en los cuales una variable se alimenta continuamente a sí misma para contribuir a su crecimiento o colapso. Por otro, lado los ciclos de retroalimentación negativa se caracterizan por un comportamiento dirigido u orientado hacia una meta.

Los ciclos de retroalimentación positiva se dividen a su vez en ciclos de retroalimentación positiva virtuosos y ciclos de retroalimentación positiva viciosos. Los ciclos de retroalimentación positiva viciosos se dan cuando un empeoramiento en un elemento de la cadena causal, ocasiona una mayor degradación del elemento, mientras que, en los ciclos de retroalimentación positiva virtuosos, un cambio positivo en un elemento del sistema adiciona un mejoramiento en él.

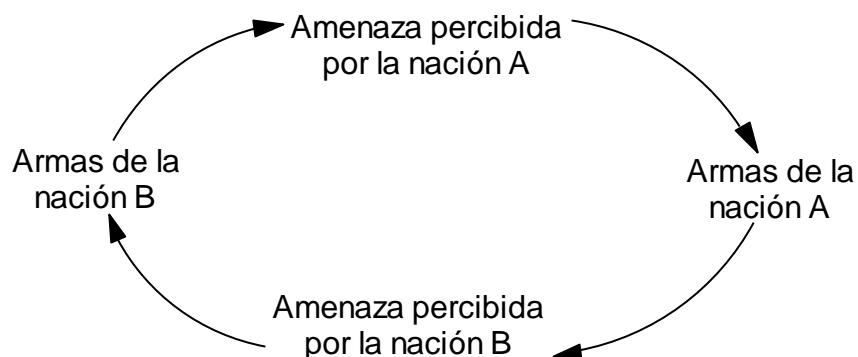
Con relación a los ciclos de retroalimentación negativa, podemos identificar en ellos algunos componentes básicos con su tipo de comportamiento y son los siguientes:

- El estado deseado o meta.
- La discrepancia.
- La acción.
- El estado del sistema o nivel.

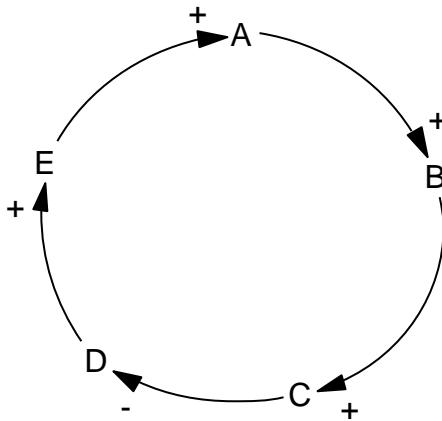
Todas estas variables definen un comportamiento cuya única finalidad es la de alcanzar una meta preestablecida la cual es controlada por algunos de estos elementos. En el siguiente capítulo se presenta la aplicación de esta teoría en un modelo conceptual conocido como *diagramas de ciclos causales* o DCC, desarrollado por el Profesor Jay Forrester en MIT.

4.11 Ejercicios

1. Coloque la polaridad de las causa-efectos e indique que tipo de ciclo se trata.



2. Explique la diferencia entre los Ciclos de Retroalimentación Positiva y Ciclos de Retroalimentación Negativa.
3. Determine la polaridad del siguiente ciclo causal.



4. Determine la diferencia que existe entre los ciclos de retroalimentación positiva virtuosa y retroalimentación positiva viciosa.
 5. Defina en sus propias palabras el método de colapso exponencial.
 6. En la Figura 4.1 de este Capítulo se sugiere un Sistema de Agua del Planeta Tierra. Coloque las respectivas polaridades de los Ciclos.
 7. Diseñe el Ciclo Causal del Nitrógeno.
 8. Mencione la importancia de los ciclos causales.
 9. Defina los términos frontera, retroalimentación, causa y efecto.
 10. Dé ejemplos de retroalimentación positiva y negativa.
-

Capítulo 5

Diagramas de Ciclos Causales

Capítulo 5

Diagramas de Ciclos Causales

Los objetivos de este capítulo son:

- Conocer todos los conceptos relacionados a los diagramas de ciclos causales.
- Aplicar estos conceptos en ejemplos simples.

5.1 Introducción

La teoría de la dinámica de sistemas enfoca la estructura y comportamiento del sistema, desde el punto de vista de sus flujos. Los diagramas de ciclos causales (DCC's) identifican los principales ciclos de retroalimentación sin distinguir entre la naturaleza de las variables que están interconectadas. En primera instancia al modelar identificamos todas las variables cualitativas y cuantitativas que tiene el problema bajo estudio. Seguidamente se identifican todas las causas-efectos, así como la polaridad de cada una de ellas. Posteriormente se procede a integrar todas las causas-efectos formando todos los ciclos que tenga el problema de la vida real. Al final, el modelador posee el primer borrador del diagrama que debe ser revisado cuidadosamente para cerciorarse que se ajusta al problema. Estos pasos parecen sencillos, pero requieren de mucho esfuerzo y cuidado para poder lograr un modelo muy cercano a la realidad.

Los DCC's juegan cuatro papeles:

- Presentan todas las causas-efectos debidamente acopladas y formando todos los ciclos cerrados que el problema tiene.
- Presentan todas las variables cualitativas y cuantitativas que tiene el problema.
- Sirven como bosquejos preliminares de hipótesis de causas-efectos.
- Simplifican la ilustración del problema a modelar.

Los DCC's permiten al analista comunicar rápidamente la proporción estructural de su modelo. En otras palabras, alienta al modelador a conceptualizar los sistemas reales en términos de ciclos de retroalimentación. Adicionalmente podemos agregar que los DCC's también sirven como instrumentos de comunicación convenientes; aunque presentan algunos defectos¹.

¹ Estos se discuten más adelante.

5.2 Definición

Un diagrama de ciclo causal (DCC) es un conjunto de **m** causas y efectos, (variables que son causa o efecto) donde las causas-efectos forman uno o varios ciclos causales. Cada ciclo se conoce como ciclo de retroalimentación y puede ser negativo o positivo. El diagrama también tiene una polaridad positiva o negativa.

5.3 Polaridad de los Ciclos Causales

Es importante que cada ciclo causal del diagrama tenga su correspondiente polaridad. Además de eso, es importante establecer la polaridad del diagrama como un todo. Esta polaridad está dada por el ciclo más externo de todo el diagrama.

Reglas:

- Cuando un ciclo de retroalimentación responde a un cambio de variables que se opone a la perturbación original o para alcanzar metas, el ciclo es negativo.
- Cuando el ciclo responde reforzando la perturbación original, el ciclo es positivo o creciente.

La misma regla (desde otro ángulo):

Sumar el número de signos negativos del ciclo:

Si es impar	→	negativo
Si es par	→	positivo

Ejemplo 1:

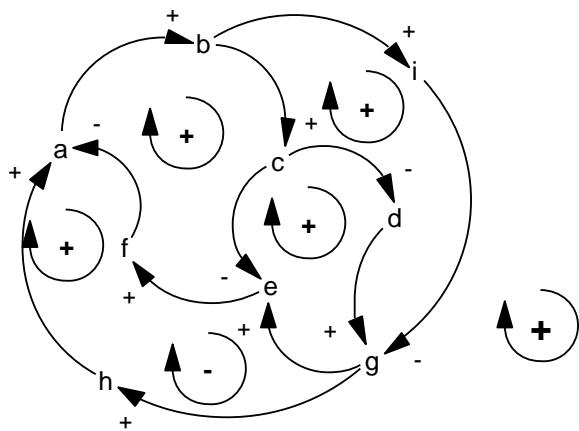


Figura 5.1: Ejemplo de Diagrama de Ciclo Causal

Note que el diagrama contiene:

- 9 variables
- 12 causas-efectos
- 5 ciclos cerrados con sus respectivas polaridades:
- a-b-c-e-f-a (Polaridad +)
- a-b-c-d-g-h-a (Polaridad -)
- a-b-i-g-h-a (Polaridad +)
- a-b-c-d-g-e-f-a (Polaridad +)
- a-b-i-g-e-f-a (Polaridad +)
- 1 polaridad global. (Positiva +)

Ejemplo 2:

El siguiente ejemplo cuenta con:

- 15 variables internas
- 2 variables externas, (Temperatura del Ambiente, Calor)
- 23 causas-efectos
- 7 ciclos Cerrados con sus respectivas polaridades:
- 1 polaridad global.

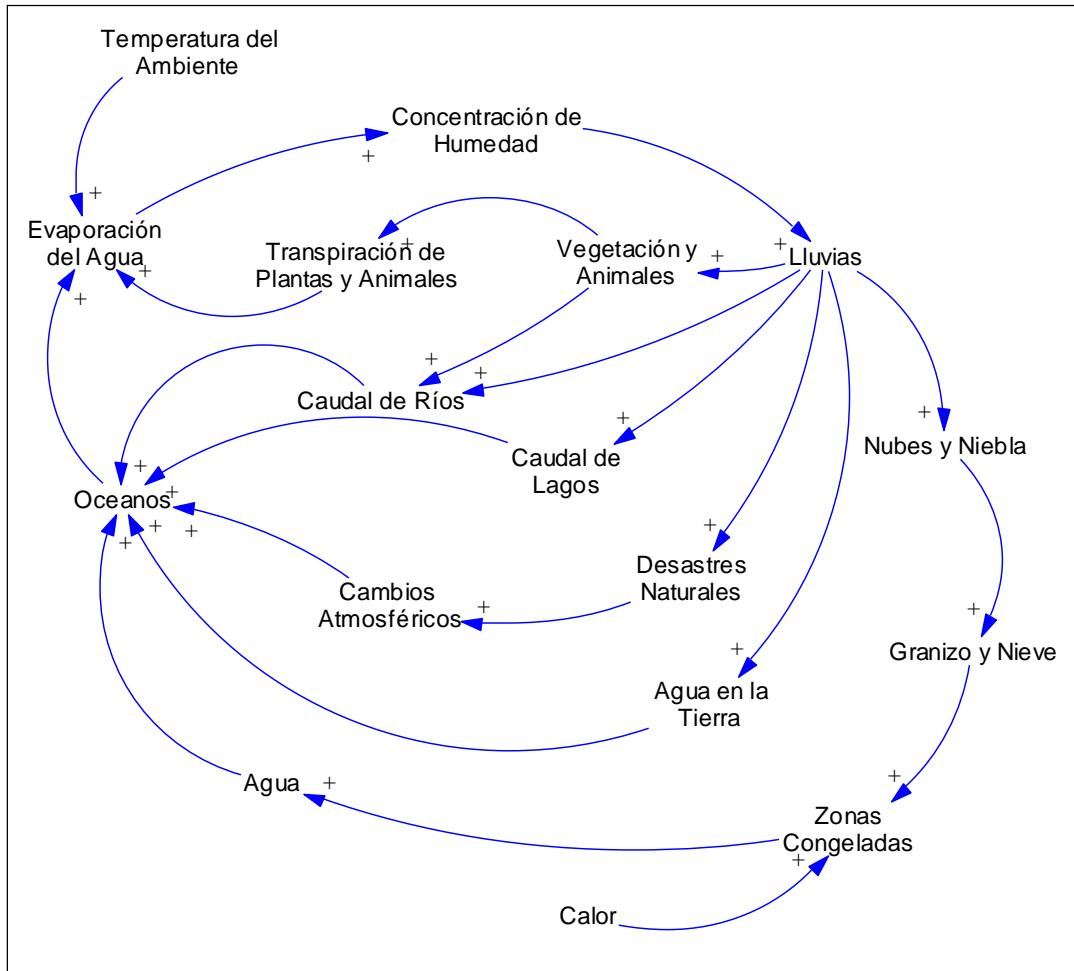


Figura 5.2: Ejemplo de Diagrama del Ciclo del Agua

5.4 Arquitectura de los Ciclos Causales Positivos CC⁺ y Ciclos Causales Negativos CC⁻

5.4.1 Ciclo de Retroalimentación Positiva CC+

Si alguna variable dentro de un ciclo de retroalimentación comienza a incrementar, entonces el efecto “bola de nieve” se presenta y la cantidad de la variable comienza a crecer y, por ende, todas las variables del ciclo.

Bola de Nieve: Su efecto está relacionado con el crecimiento de ciertas ideas, o modos de acción. Este comportamiento se asemeja con el crecimiento de una bola de nieve, al rodar ésta hacia abajo en una montaña; como en su caída, recoge nieve, su masa y circunferencia se incrementan lo cual causa el crecimiento rápido de ésta.

Los círculos viciosos o virtuosos son sinónimos de retroalimentación positiva. En un círculo vicioso, el empeoramiento de un elemento en la cadena causal ocasiona además una degradación del elemento. Por otro, lado, en el círculo virtuoso, un cambio positivo en un elemento del sistema, nos adiciona un mejoramiento en él. La viciosidad o virtuosidad de un sistema de retroalimentación positiva, depende si los elementos del ciclo se deterioran o se mejoran mutuamente.

Ejemplo 1:

La espiral inflacionaria del salario-precio nos provee de un ejemplo clásico de círculo vicioso. Los trabajadores, inquietos por los altos precios, demandan salarios superiores. Los fabricantes aumentan sus precios para cubrir el costo del trabajo. Los trabajadores perciben que los precios suben otra vez y demandan salarios superiores para cubrir estos incrementos. La siguiente figura nos muestra el diagrama de ciclo causal de la espiral. Cada variable en el ciclo exhibe el mismo comportamiento. Es decir, el comportamiento de las variables es creciente. Por ejemplo, el aumento de la variable salario ocasiona el aumento de la presión de precios.

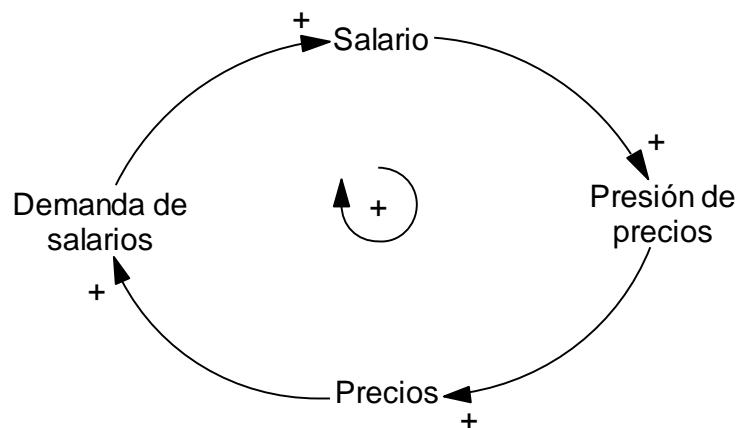


Figura 5.3: Salarios y Precios

Ejemplo 2: ¿Cuándo crece, cuándo decrece?

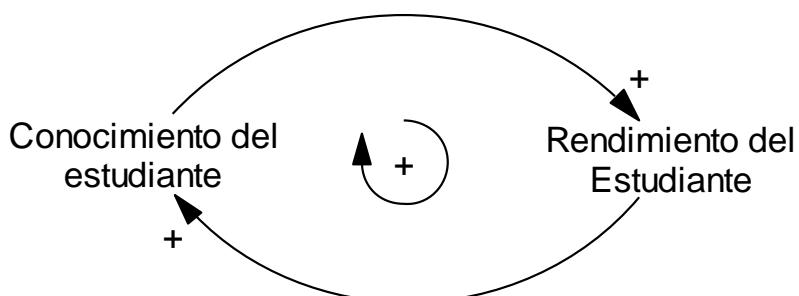
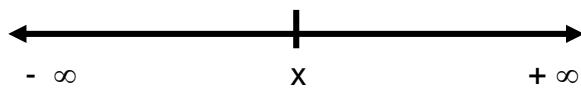


Figura 5.4: Conocimiento y Rendimiento

La respuesta dependerá del nivel de conocimiento y de rendimiento del estudiante antes de que se comience a ejecutar el ciclo. Los niveles de las variables se les conoce como *puntos de equilibrio*. Un ciclo de retroalimentación positivo tiende a crecer hacia infinito o a decrecer hacia menos infinito. Su comportamiento es tender hacia cada vez más o menos.



Determinar los puntos de equilibrio es difícil en la mayoría de los casos. También es cierto, que la mayoría de los ciclos de retroalimentación positivos se estabilizan por medio de otros ciclos que no han sido considerados en análisis iniciales.

5.4.2 Ciclo de Retroalimentación Negativo: CC-

El comportamiento de un ciclo de retroalimentación negativo se describe también utilizando el punto de equilibrio. El ciclo de retroalimentación positivo tiende hacia el punto de equilibrio, por eso es que es estable.

La retroalimentación negativa es caracterizada por una meta directa o un comportamiento orientado hacia una meta. Términos tales como auto-gobernación, auto-regulación, auto-equilibrio, homeostático, o adaptable, todos implican la presencia de una meta, y definen sistemas de retroalimentación negativa.

El controlador termostático del calentador es un sistema común de auto-gobernación. El sistema de termostato procura mantener la temperatura del cuarto que ha sido preseleccionada. La unidad que toma las decisiones, el termostato, siente una disparidad entre la temperatura deseada y la actual de la habitación, y activa la unidad de calentamiento. El calor adicional aumenta eventualmente la temperatura del cuarto al nivel deseado. Luego el termostato automáticamente corta la calefacción.

Ejemplo: Sistema de Calentamiento del Termostato

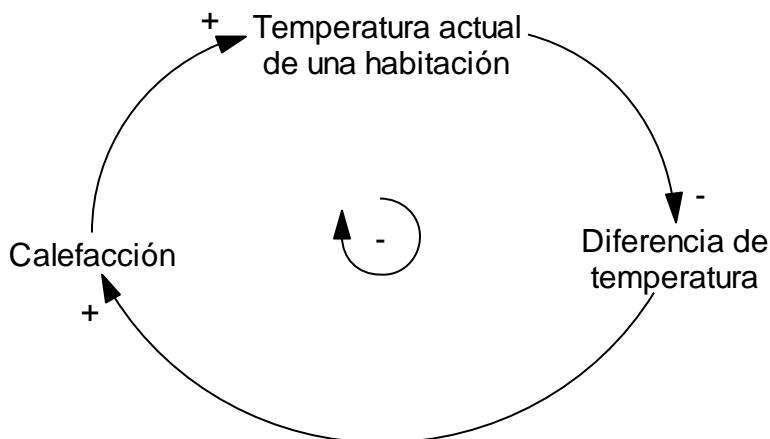


Figura 5.5: Ejemplo de Calefacción

5.5 Ejemplo 1: Ciclo del Trabajo Urbano²

La mejor forma de entender los diagramas de ciclos causales es presentando un caso de la vida real. A continuación, se presentan las cinco hipótesis causales sencillas del comportamiento laboral urbano, donde la Disponibilidad de Trabajo es igual a DT, Inmigración igual a I, Trabajo a T y finalmente Población Laboral es igual a PL:

- La disponibilidad de trabajos DT atrae inmigrantes I a la ciudad.
- Los nuevos habitantes de la ciudad expanden la población laboral PL.
- La población laboral absorbe los trabajos disponibles, disminuyendo la cantidad de disponibilidad de trabajos DT.
- A largo plazo, la población laboral también crea una demanda de mercancía adicional, servicios humanos y facilidades, y nuevamente se incrementa el número total de trabajos en el área afectada.
- Más trabajos T incrementan la disponibilidad de trabajos DT.

El problema consta de cuatro (4) variables. Las definiciones de las variables respectivas son las siguientes:

Variables (nombre largo)	Variables (nombre corto)	Definición
Trabajo	T	Número de vacantes y cantidad de trabajos que tienen las personas en el área de estudio.
Disponibilidad de Trabajo	DT	Número de vacantes en el área de estudio ³ .

² Forrester Jay, "Systems Dynamics".

³ T-DT nos da la población laboral existente PL

Inmigración	I	Inmigración neta hacia el área urbana o de estudio.
Población Laboral	PL	Población laboral total del área en estudio.

Para la simplificación del diagrama de este ejemplo ignoramos factores como: El tipo de trabajo, las características demográficas de migración, entre otros. El desarrollo paso a paso del diagrama ilustrará la mecánica de diagramación de ciclos. Para la estructura de los sistemas en diagramas de ciclos causales y la identificación de la polaridad de cada ciclo, debemos establecer primero la relación de paridad entre las variables pertinentes.

Se asume que el número de trabajos disponibles regula el flujo de personas en el área urbana. Por ejemplo, un incremento en la disponibilidad del trabajo causa un incremento en la migración hacia el área. Una disminución en los empleos disponibles tiene un efecto opuesto. La representación causal para esta proposición es mostrada en la figura 5.6.

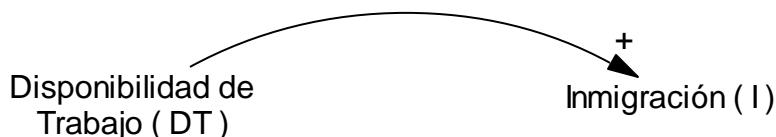


Figura 5.6: Causa-Efecto Positiva de Disponibilidad de Trabajo - Inmigración

La flecha indica la dirección de influencia; el signo (positivo o negativo), el tipo de influencia⁴. Un incremento (decremento) en la disponibilidad de trabajo DT debería producir un incremento (decremento) en la inmigración I. La relación tiene signo positivo que significa el carácter “positivo” de la unión. Un cambio en una variable genera un cambio en la misma dirección en la segunda variable relativa a su valor anterior. Para aplicar esta definición o prueba, debemos considerar solo pares adyacentes de variables. La próxima causa-efecto involucra las variables inmigración I y población laboral PL. Se representa la relación en la misma forma que la relación DT-I. La figura 5.7 muestra la relación I-PL. Un incremento en la tasa I de inmigración, atrae trabajadores adicionales, quienes incrementan la población laboral residente.

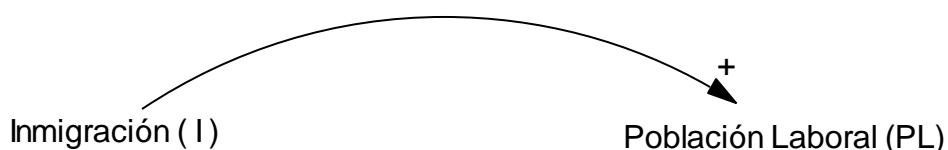


Figura 5.8: Unión Positiva Inmigración-Población Laboral

⁴ Tal como se explicó en el capítulo anterior

En la siguiente causa efecto surge una relación negativa denotada por un signo menos y ocurre cuando un cambio en una variable tipo causa produce un cambio en la dirección contraria en la segunda variable tipo efecto. La figura 5.8 ilustra una relación negativa.

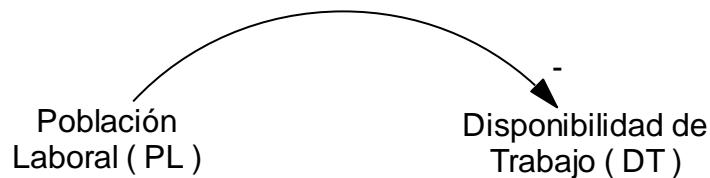


Figura 5.8: Causa-Efecto Negativa de Población Laboral-Disponibilidad de Trabajo

La figura 5.8 asume que un incremento en la población laboral residente eventualmente decrementará el número de trabajos disponibles en el área urbana.

Los nuevos trabajadores en la ciudad ocuparán los trabajos disponibles, reduciendo así la disponibilidad de trabajos DT. Si la población laboral disminuyera, asumimos que habría más trabajos disponibles. En otras palabras, incrementando o decrementando la población laboral se produce un cambio opuesto en DT. Examinando la relación entre la población laboral PL y la Disponibilidad de Trabajos DT, ignoramos el hecho de que los trabajos también cambian. Ahora bien, un aumento de la población laboral PL causa que el número de trabajos T también aumenten.

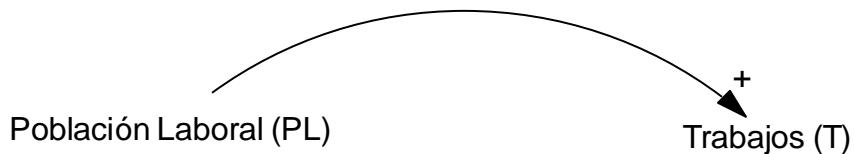


Figura 5.10: Causa-Efecto Positiva de Población Laboral-Trabajos

Un cambio en el número total de trabajos en el área, por supuesto, puede cambiar la disponibilidad de trabajo DT. La figura 5.10 ilustra esa simple relación positiva. Un incremento en el número total de trabajos T asumiendo que todo lo demás (incluyendo la fuerza laboral) permanece constante, puede incrementar la disponibilidad de trabajo DT. La suposición que todas las otras variables relacionadas permanecen constantes durante la determinación de la polaridad causal es central al análisis causal.

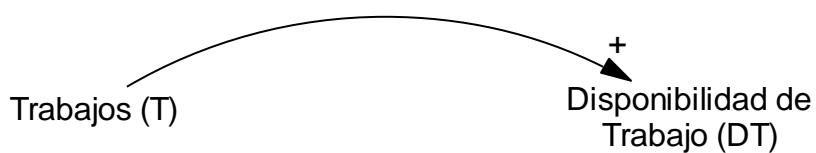
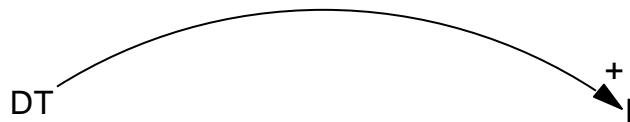


Figura 5.11: Causa-Efecto Positiva de Trabajos-Disponibilidad de Trabajo

Diccionario de Causas-Efectos:

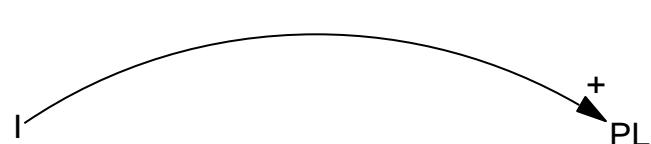
Estas cinco causas-efectos se resumen en un *cuadro de causas-efectos del problema*.

1:



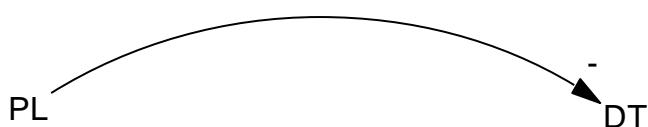
Asumimos que el número de trabajos disponibles, regula el flujo de personas al área en estudio. Un aumento en los trabajos disponibles causa un incremento en la inmigración. Una disminución en los empleos disponibles tiene efecto opuesto.

2:



Un incremento en la tasa de inmigración, incrementa el número de trabajadores los cuales aumentan la población laboral.

3:



Un aumento en la población laboral disminuye la disponibilidad de trabajos en el área en estudio. El decremento causa un efecto opuesto.

4:



Un incremento en la población laboral incrementa los trabajos en el área en estudio. Lo opuesto tiene el mismo efecto.

5:



Un cambio positivo en el número total de trabajadores en el área de estudio por supuesto que cambia positivamente la disponibilidad de trabajo.

Cuadro de Causas-Efectos:

Hipótesis	Definición
1	La Disponibilidad de Trabajo DT atrae Inmigrantes I al área en estudio.
2	Los nuevos habitantes del área en estudio expanden la Población Laboral PL.
3	La Población Laboral PL absorbe los Trabajos disponibles y por ende disminuye la Disponibilidad de Trabajo DT.
4	A largo plazo, la Población Laboral PL también crea una demanda de bienes servicios y facilidades adicionales, incrementando así el número total de trabajos T en el área en estudio.
5	Más Trabajos T incrementan la Disponibilidad de Trabajos DT.

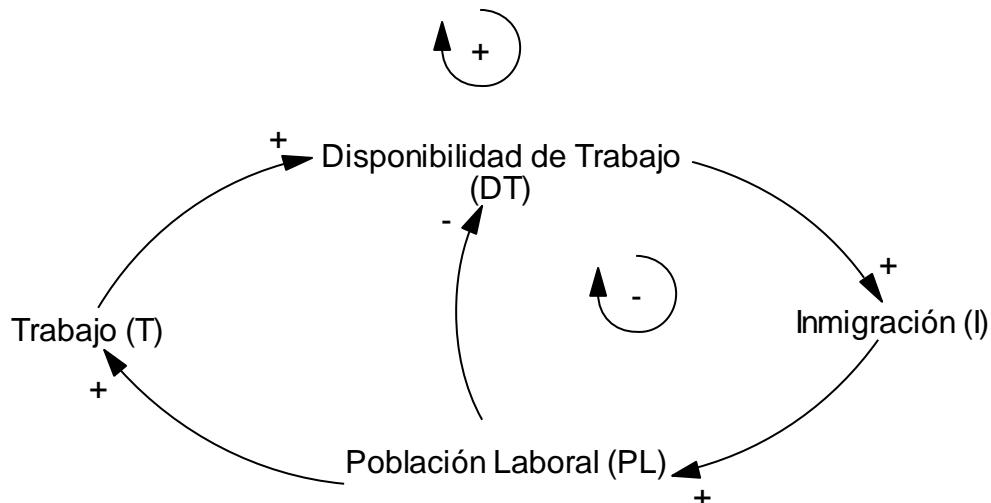


Figura 5.11: Diagrama de Ciclos Causales del Trabajo Urbano

Finalmente se acoplan todas las causas-efectos y se forma el diagrama de ciclo causal DCC para el problema simplificado del Trabajo Urbano (véase la figura 5.11).

Note que el diagrama tiene:

- 4 variables
- 5 causas-efectos
- 2 ciclos cerrados (I-PL-DT-I; I-PL-T-DT-I) con sus respectivas polaridades y
- 1 polaridad global.

5.6 Ejemplo 2: Acondicionador de Aire

Cabe mencionar que los seres humanos tienen una debilidad inherente: necesitan sentirse cómodos. Quieren vivir en un ambiente que no sea ni caliente, ni frío, ni muy húmedo, ni muy seco. Sin embargo, la comodidad no se obtiene con facilidad ya que las necesidades del cuerpo humano y el clima suelen no ser muy compatibles. Para alcanzar la comodidad es necesario luchar constantemente contra los factores que causan la incomodidad, como temperaturas mayores o menores y humedades altas o bajas.

5.6.1 Definición

El acondicionador de aire es un sistema compuesto por diferentes elementos mecánicos, electrónicos y eléctricos que interactúan entre sí para realizar un trabajo, lo cual permite reducir la temperatura a través de la remoción de calor de espacios, para mantener la confortabilidad de los mismos.

Acondicionar el aire es regular las condiciones ambientales, ya sea calefaccionando o refrigerando, de manera tal que la temperatura, humedad y grado de pureza del aire sean las adecuadas. Por ejemplo, se acondiciona el aire para realizar algunos procesos industriales, en las salas de terapia intensiva de los hospitales, en los centros de cómputos, etc.

El objetivo del acondicionador de aire, es el de encargarse de producir frío o calor y de impulsar el aire tratado a la vivienda o local.

5.6.2 ¿Cómo funciona un Acondicionador de Aire?

El acondicionador de aire toma aire del interior de una recámara pasando por tubos que están a baja temperatura estos están enfriados por medio de un líquido que a su vez se enfriá por medio del condensador, parte del aire se devuelve a una temperatura menor y parte sale expulsada por el panel trasero del aparato, el termómetro está en el panel frontal para que cuando pase el aire calcule la temperatura a la que está el ambiente dentro de la recámara, y así regulando que tan frío y que tanto debe trabajar el compresor y el condensador.

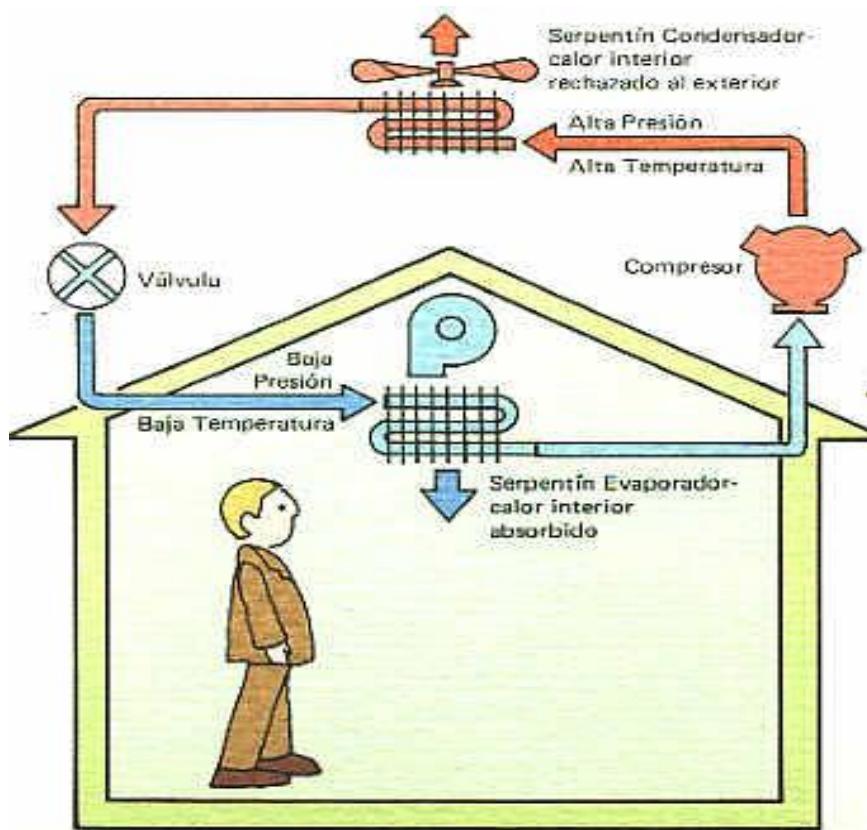


Figura 5.12: Funcionamiento del Acondicionador de Aire

En la figura se muestra un acondicionador de ventana tradicional, funcionando en condiciones típicas.

5.6.3 ¿Qué Parámetros controla el Equipo de Acondicionador de Aire?

Temperatura

Es evidente, pero también muy importante: el acondicionador de aire elimina el calor del aire, por lo que el termómetro señala una temperatura más fresca.

Humedad

El aire de nuestro entorno natural contiene humedad en forma de vapor, y el nivel saludable de humedad relativa del aire se sitúa entre el 40-60%. La capacidad del aire de contener vapor de agua depende sobremanera de su temperatura: el aire cálido puede atrapar más humedad que el fresco. Como el aire cálido atraviesa las baterías frías del sistema de acondicionador de aire, el exceso de humedad desaparece del aire al condensarse en las baterías, recogiéndose en forma de líquido en el sistema.

Pureza del Aire

El aire está repleto de partículas diminutas: polvo, polen, humo, contaminación, bacterias, esporas y otras muchas impurezas. El sistema de acondicionador de aire está dotado de un filtro que retiene tales partículas y devuelve al interior un aire depurado y más saludable.

Circulación del Aire

Todos estos efectos beneficiosos no se repartirían de manera uniforme por los recintos si el sistema no renovase el aire de forma óptima. El ventilador y las rejillas del sistema deben ajustarse de forma que hagan circular el aire de manera uniforme en el lugar de instalación, pero sin expulsarlo directamente hacia las personas situadas junto a ellos.

5.6.4 Componentes del Acondicionador de Aire

Compresor

Comúnmente denominado el corazón del sistema, como su nombre lo indica, comprime el gas refrigerante tomando para ello potencia del motor mediante una transmisión de correa. Los sistemas de acondicionador de aire están divididos en dos lados: el lado de alta presión y el lado de baja presión; también denominados descarga y succión respectivamente. La entrada del compresor toma el gas refrigerante de la salida del evaporador, y en algunos casos lo hace del acumulador, para comprimirlo y enviarlo al condensador.

Condensador

Aquí es donde ocurre la disipación del calor. El condensador tiene gran parecido con el radiador debido a que ambos cumplen la misma función. El condensador está diseñado para disipar calor, y normalmente está localizado frente al radiador, pero a veces, debido al diseño aerodinámico de la carrocería del vehículo, se coloca en otro lugar. El condensador debe tener un buen flujo de aire siempre que el sistema esté en funcionamiento. Dentro del condensador, el gas refrigerante proveniente del compresor, que se encuentra caliente, es enfriado; durante el enfriamiento, el gas se condensa para convertirse en líquido a alta presión.

Evaporador

El evaporador está localizado dentro del vehículo, y sirve para absorber tanto el calor como el exceso de humedad dentro del mismo. En el evaporador el aire caliente pasa a través de las aletas de aluminio unidas a los tubos; y el exceso de humedad se condensa en las mismas, y el sucio y polvo que lleva el aire se adhiere a su vez a la superficie mojada de las aletas, luego el agua es drenada hacia el exterior.

En el proceso de evaporación el refrigerante absorbe grandes cantidades de calor, el cual es llevado por el refrigerante fuera del interior del recinto. Existen otros componentes de los sistemas de acondicionador de aire que trabajan en conjunto con el evaporador, puesto que deben existir controles para mantener la presión baja, y la temperatura, puesto que, si ésta disminuye por debajo del valor mencionado anteriormente, el agua producto de la condensación del exceso de humedad no solo se condensará, sino que se congelará alrededor de los tubos del evaporador, y esto disminuye la eficiencia de la transferencia de calor en el mismo.

Dispositivos Reguladores de Presión

La temperatura del evaporador puede ser controlada mediante la regulación del flujo y la presión del refrigerante dentro del mismo. Existen muchos dispositivos creados para tal fin. Son instrumentos de precisión, designados a regular el flujo del líquido refrigerante que entra al evaporador en la misma proporción que este líquido refrigerante se evapora dentro del evaporador. Las válvulas responden a la temperatura del gas refrigerante que sale del evaporador y a la presión dentro del evaporador.

Acumulador

Los acumuladores normalmente son utilizados en sistemas que utilizan un tubo orificio, y están conectados a la salida del evaporador, en donde se almacena el exceso de líquido que no se evaporó, debido a que si este líquido pasa al compresor éste se puede dañar; aunque ésta es su función principal, el acumulador también sirve para eliminar la humedad y las impurezas.

Refrigerantes

Desde mediados de la década del 80, los sistemas de refrigeración comerciales han pasado por un proceso de transición, desde usar compuestos refrigerantes que reducen el ozono, incluyendo los clorofluorocarburos (CFC), a compuestos de baja o ninguna reducción de ozono, tal como los hidroclorofluorocarburos (HCFC) y los hidrofluorocarburos (HFC). El amoníaco y la absorción se usan en menor medida.

Termostato

Un termostato percibe su entorno y realiza acciones sobre él (subir / bajar la temperatura) teniendo en cuenta sus consecuencias (intenta que se retire calor suficiente para compensar la diferencia de temperatura); también es autónomo, no necesita que nadie le diga cuándo debe hacer sus tareas; y se trata de un elemento integrado en su entorno reaccionando dinámicamente a los cambios que suceden.

Los termostatos son normalmente placas metálicas que mediante la elevación o la disminución de la temperatura se expanden o se contraen térmicamente. Cada material tiene un coeficiente de expansión o de dilatación térmica.

Los termostatos cuentan con un bulbo que mide la temperatura en el evaporador en un momento dado; esta temperatura es transmitida a las placas bi-métálicas del termostato, cuando esta temperatura llega al punto de consigna del selector del termostato, se produce una separación entre las placas bi-métálicas y esto se traduce en un circuito abierto, lo que significa que no habrá flujo de corriente hacia el compresor, por lo cual el ciclo de refrigeración se mantendrá detenido, hasta que se produzca una diferencia de temperatura que haga que estas placas se vuelvan a unir, lo que hace que inicie el sistema de refrigeración.

En la actualidad existen termostatos que utilizan censores para medir la temperatura del ambiente.

5.6.5 Descripción del Acondicionador de Aire como Sistema

ENTIDAD	ATRIBUTOS	ACTIVIDADES
Compresor	Eficiencia volumétrica, desplazamiento del compresor, eficiencia Isentrópica, potencia eléctrica aplicada, potencia eléctrica producida y mecánica, relación de compresión, velocidades angulares de operación y máximas permitida (rpm), tensión de las bandas o correas, tiempo de trabajo.	Elevar la presión y la temperatura del refrigerante y mantener el flujo másico del refrigerante.
Condensador	Diámetro de la tubería, conductividad térmica del material, volumen de refrigerante que acumula, área de contacto con el fluido con el que intercambiará calor, eficiencia del diseño para la transferencia de calor.	Enfriar el refrigerante y lo condensa.
Evaporador	Diámetro de la tubería, conductividad térmica del material, volumen de refrigerante que acumula, área de contacto con el fluido con el que intercambiará calor, eficiencia del diseño para la transferencia de calor.	Evapora el refrigerante que entra a baja temperatura y en él se da el fenómeno de absorción de un espacio refrigerado
Dispositivos reguladores de presión (válvula de expansión)	Conexión, medida, ecualizador refrigerante, ángulo, rango.	Estrangula el refrigerante procedente del evaporador disminuyendo su temperatura y presión.

Refrigerante	Estabilidad química, momento del dipolo, punto de la burbuja, punto y presión de condensación, punto y presión de evaporación deslizamiento de la temperatura, flujo másico, viscosidad, presión, calor específico (C_p y C_v).	Absorber el calor del espacio refrigerado y este calor lo vierte a un medio exterior por medio del condensador.
Termostato	Rango de control de temperatura, tiempo de respuesta, tipo de control (todo o nada, proporcional, PID, etc.)	Controla los rangos de temperatura del espacio refrigerado.

Ambiente

El acondicionador de aire se encuentra rodeado (para el caso del salón) de sillas, paredes, ventanas, personas.

Relaciones

A continuación, se presentan las relaciones básicas de los componentes del acondicionador de aire:

El rango de control de temperatura del termostato se relaciona con el tiempo de trabajo del compresor, ya que el termostato es el que controla el tiempo en que el compresor se va a mantener funcionando.

Las velocidades angulares de operación se relacionan con la rapidez del flujo de refrigerante a través de los dispositivos, o sea, el flujo másico del refrigerante.

El diámetro de la tubería del evaporador se relaciona con la viscosidad del refrigerante.

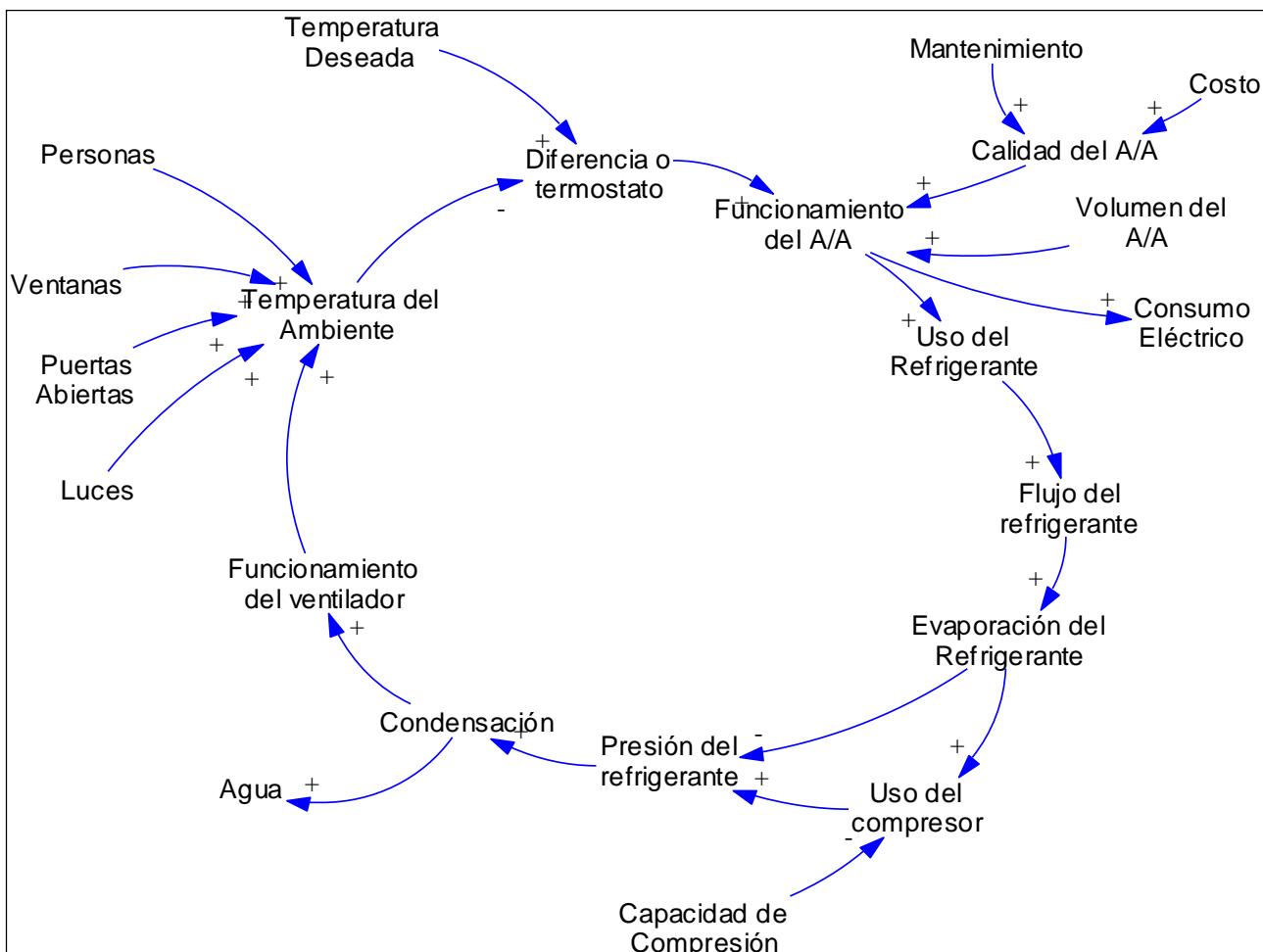
El evaporador y condensador por medio de la conductividad térmica se relaciona con el refrigerante por medio de sus calores específicos o por la capacidad de transferir calor.

El compresor por medio de la relación de compresión se relaciona con el refrigerante por el punto y presión de evaporación.

Estados

- Compresor: Encendido o Apagado.
- Evaporador: Con las Tomas Abiertas.
- Condensador: Con las Tomas Abiertas.
- Válvula de Expansión: Abierto o Cerrada.
- Gas Refrigerante: Estático o en Circulación
- Termostato: Relay Abierto o Cerrado.

5.6.6 Diagrama de Ciclo Causal del Acondicionador de Aire



5.7 Diversidad de Comportamiento⁵

La actividad dentro de los ciclos de retroalimentación "negativa" puede mostrar desde un tranquilo logro de la meta que busca el ciclo, hasta fuertes fluctuaciones. Los ciclos de retroalimentación "positiva" muestran decadencia o crecimiento (véase la figura 5.12).

La curva A representa el más sencillo de los sistemas de retroalimentación, donde la variable tiene un crecimiento cuya tasa va decreciendo hacia un valor final. El valor de la variable mostrada por la **curva A**, es cero al comienzo, donde el comienzo está definido en el tiempo cero.

⁵ Forrester Jay, "Principles of Systems"

Los valores observados serían:

Tiempo:	0	1	2	---
Variable:	0	1.3	2.05	3

Y se sigue aproximando sin llegar realmente al último valor que sería 3. Tal aproximación simple al equilibrio podría representar el aumento de los empleos en una compañía hasta un nivel permisible. O la curva puede representar la información que da un aparente nivel (esto es, el estado o la condición) de un sistema y cuyo entendimiento crece hasta el valor real. En todo caso, el cambio es más rápido al principio y se aproxima más y más lentamente hacia el valor final a medida que la discrepancia disminuye entre el valor presente y el final.

La **curva B** es más complicada para aproximarse al valor final (aquí de nuevo, es el valor de 3), donde el sistema sobrepasa el valor final luego cae por debajo tratando de recuperarse. Tal comportamiento se debe a excesivos retrasos en el ciclo de retroalimentación o también porque se puede ver un esfuerzo muy rápido por tratar de conseguir una nivelación. Tal fluctuación puede observarse en la velocidad variable de un motor con gobernador automático defectuoso, en la subida y bajada de la producción industrial como se ve en el ciclo económico, en la fluctuación de precios de los artículos para abastecimiento, prestando la debida atención a la demanda.

La **curva C** muestra un crecimiento, donde cada intervalo de tiempo, sufre el mismo crecimiento fraccional en la variable. En la ilustración, el valor vertical se duplica en cada unidad de tiempo. Este crecimiento “exponencial” puede verse en la división de las células, en el crecimiento de las ventas donde los vendedores venden para que se puedan contratar más vendedores, en la reacción en cadena de una explosión atómica y en la multiplicación de los conejos.

La **curva D** muestra una sección inicial de crecimiento exponencial seguido por una salida nivelada. La curva D tiene un comportamiento inicial igual a la curva C; después tiene un comportamiento final ajustado, es decir, un comportamiento final como la curva B. En algunos casos el comportamiento final tiende a ser como la curva A.

Tal comportamiento se puede apreciar en el crecimiento de un animal, el cual, al principio, aumenta rápidamente y luego se aproxima, despacio, a su tamaño final. Este tipo de crecimiento, provee una forma para un continuo balance, puede representarse también en la población de conejos, que crece rápidamente hasta el punto donde ya no hay comida disponible para todos y no pueden mantenerse más conejos. La **curva D** se ve en la actividad nuclear de una planta atómica donde la tasa de fisión aumenta el nivel de operaciones y es controlado por el sistema de central. También representa el rápido crecimiento de un producto que luego se estanca ya que la demanda del mercado ha sido satisfecha, o debido a que la capacidad de producción ha sido alcanzada, o porque su calidad ha declinado.

Otros Ejemplos:

- La cantidad de calor en un termómetro determina la temperatura ambiental, a mayor calor mayor temperatura. La rata a la que el calor fluye depende de la diferencia de temperatura entre el termómetro y su alrededor. A medida que la temperatura de un termómetro se aproxima a la temperatura ambiental, el calor disminuye gradualmente. La **Curva A** describe como la temperatura de un termómetro cambia con el tiempo después de ser sumergido en un líquido caliente.
- La posición de un péndulo que es desplazado y se balancea, es descrita mejor con la **Curva B**.
- Un estudiante, que aprende una nueva materia, encuentra que, al incrementarse su conocimiento, le será, a su vez más fácil adquirir más conocimiento. Él aprende a pasos rápidos hasta que comienza a aproximarse a su capacidad máxima en la materia, entonces pierde interés o se desanima porque no está haciendo el mismo progreso rápido, así como al comienzo y puede reducir su motivación. La **Curva D** describe este proceso.
- En la industrialización de una sociedad, el capital en equipo puede ser usado para producir, más capital en equipo. La **Curva C** describe la cantidad de capital en equipo versus el tiempo: El resultado acumulado de acciones pasadas (el capital en equipo disponible) determina la acción presente (rata de construcción de más equipo).

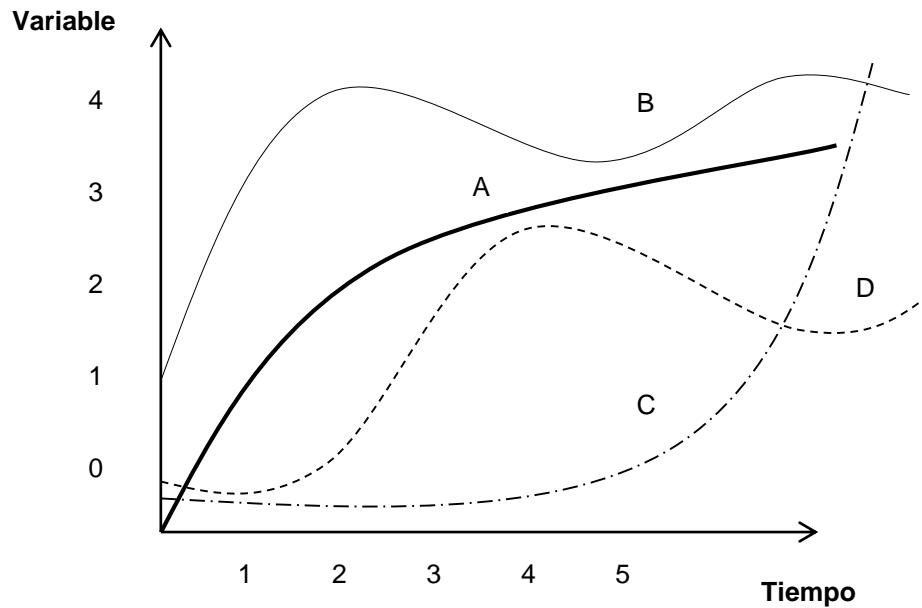


Figura 5.13 Comportamiento Dinámico

La actividad dentro de los ciclos de retroalimentación *negativa* puede mostrar desde un tranquilo logro de la meta que busca el ciclo, hasta fuertes fluctuaciones en busca de la meta. Los ciclos de retroalimentación “positiva” muestran decadencia o crecimiento. El acoplamiento “no-lineal” puede causar un cambio de dominio desde un ciclo del sistema a otro. Esta sección presenta varios sistemas de retroalimentación para ilustrar las clases de reacciones representadas en el dibujo a continuación. Estas curvas muestran el valor cambiante de algunas variables del sistema a medida que el tiempo pasa de izquierda a derecha.

5.8 Beneficios

Los diagramas de ciclos causales pueden ser útiles durante las etapas iniciales de la conceptualización de un modelo, ya que nos ayudan a identificar y organizar los componentes principales y los ciclos de retroalimentación del sistema que se está estudiando. Sin embargo, los diagramas de flujo dinámico son indispensables en las etapas posteriores de la construcción del modelo y la escritura de ecuaciones.

5.9 Desventajas

Los diagramas de ciclos causales carecen de precisión, explicación de tipos de variables (detalle de las variables tipo flujo o nivel o elementos auxiliares encontrados en diagramas de flujos⁶). Por esto los diagramas de flujos llevan errores conceptuales que no se muestran prontamente en el diagrama de ciclo causal. Por ejemplo, un diagrama de ciclo causal puede contener inadvertidamente ciclos cerrados sin niveles. Tal error se encontraría fácilmente en un diagrama de flujo.

Un diagrama de flujo aporta considerablemente más información que los diagramas de ciclo causal, acerca de la estructura del sistema y su comportamiento. Por describir ellos la de flujo pueden indicar muchas veces posibles tipos o modos de comportamiento de un sistema. Por ejemplo, el orden de un sistema limita los modos de comportamiento posibles de ese sistema. Por esto, el conocimiento del orden del sistema es esencial para pronosticar su comportamiento. Tal información está contenida explícitamente en los diagramas de flujo, pero no en los diagramas de ciclo causal.

Los diagramas de ciclo causal muchas veces no proporcionan la información necesaria, para entender el comportamiento de una estructura de ciclos de retroalimentación, tal como los procesos retrasados. Sin embargo, estos procesos forman ciclos de retroalimentación negativos que podrían tener un efecto considerable en el comportamiento del sistema.

⁶ T-DT → Esta funcionalidad se presenta en los diagramas de flujos dinámicos

5.10 ¿Cómo se construye un Diagrama de Ciclos Causales?

Paso 1: Análisis del Problema

En primera instancia se recolectan trabajos de investigación sobre el problema en estudio que le permitan al ingeniero de sistemas comprender exhaustivamente el problema.

Paso 2: Listado de Variables

De las investigaciones, estudios de campo, entrevistas, etc. se extraen todas las variables del problema. Un caso de nivel medio puede llegar a tener más de 100 variables.

Existen dos tipos de variables:

- Nivel: Poseen un valor cuantitativo o cualitativo.
- Flujo: Miden, a través del tiempo, la rata o movimiento de un objeto.

Por ejemplo, un tanque de reserva de agua posee 100,000 galones de agua que son enviados a tanques menores que se encuentran cercanos a un pueblo. Entre el tanque grande y los chicos existen tuberías de agua que tienen ciertos diámetros y el agua fluye a cierta velocidad, dependiendo de la diferencia de alturas entre el tanque grande y los tanques medianos, así como de si existen bombas que aceleren el agua, entre otros factores. Cada tanque tiene, en una unidad de tiempo específica, una cantidad de agua y por las tuberías fluye el agua a ciertas velocidades. Bien los contenidos de los tanques son *variables de nivel* y los flujos a través de las tuberías, son *variables de rata o flujo*. En tipo de variable se coloca si es de nivel o de flujo. La descripción debe definir explícitamente la variable sin dejar una posible mala interpretación.

Nombre Corto	Nombre Largo	Tipo	Descripción

Figura 5.14 Cuadro para listar las Variables

Paso 3: Agrupación de las Variables

Las variables se agrupan en diferentes tipos, formando grupos, tal como si elaboráramos un mapa de contenido:

- Materiales
- Equipo capital
- Dinero (financieros, económicos, contables, etc.)
- Ordenes (leyes, políticas, manuales, procedimientos, etc.)
- Personas o animales o plantas.
- Información

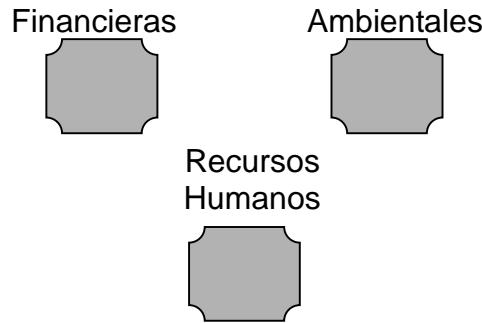


Figura 5.15: Agrupación de Variables

Es recomendable que a cada grupo de variables les asignemos un color y pueden existir otras agrupaciones. Esto depende del problema bajo estudio. Las variables que van de un grupo a otro se les conoce como *variables de acoplamiento* y siempre son variables que transmiten información; no pueden ser de ningún otro tipo.

Paso 4: Causas-Efectos

Dependiendo de los estudios científicos elaborados sobre el problema en estudio, se elabora un listado de cada causa-efecto que se encuentre científicamente sustentada. A cada causa-efecto se le señala la sustentación científica.

Causa - Efecto	Descripción

Figura 5.16: Cuadro de Causa-Efectos

Paso 5: Diagramación

Finalmente se procede a conectar todas las causas-efectos formando el primer borrador del diagrama.

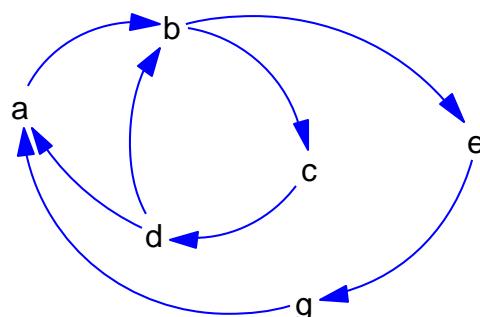


Figura 5.17: Ejemplo de Diagrama de Ciclo Causal

Paso 6: Revisión y Validación

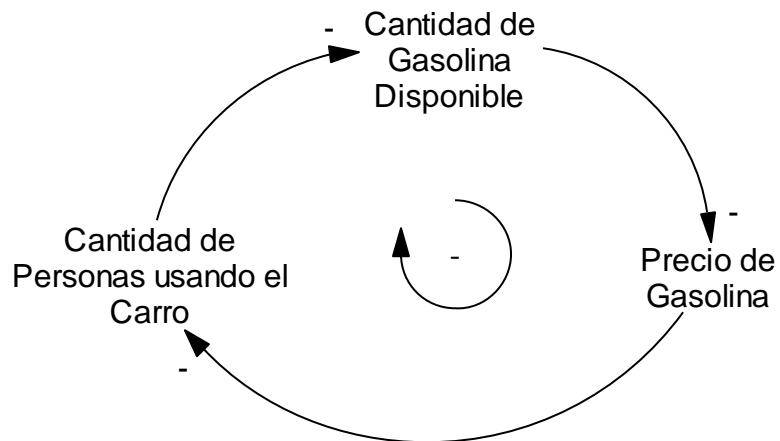
En este paso, que puede ser muy demorado, se revisa cuidadosamente el diagrama y se validan nuevamente todas las causas-efectos, así como los ciclos causales que surjan.

5.11 Resumen

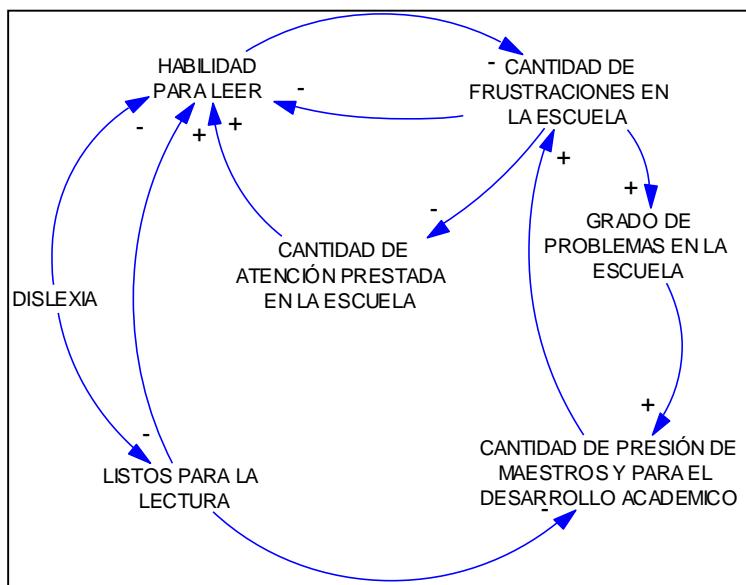
Los diagramas de ciclos causales pueden ser útiles durante las etapas iniciales de la conceptualización de un sistema, ya que ayudan a identificar y organizar los componentes principales y los ciclos de retroalimentación del sistema que se está estudiando, pero, a pesar de esto, los diagramas de ciclos causales muchas veces no proporcionan la información necesaria para entender el comportamiento de una estructura de ciclos de retroalimentación, tal como los procesos retrasados. Sin embargo, estos procesos forman ciclos de retroalimentación negativos, que podrían tener un efecto considerable en el comportamiento del sistema

5.12 Ejercicio 1

1. De acuerdo a las reglas de la polaridad identifique la polaridad de los siguientes ciclos causales:



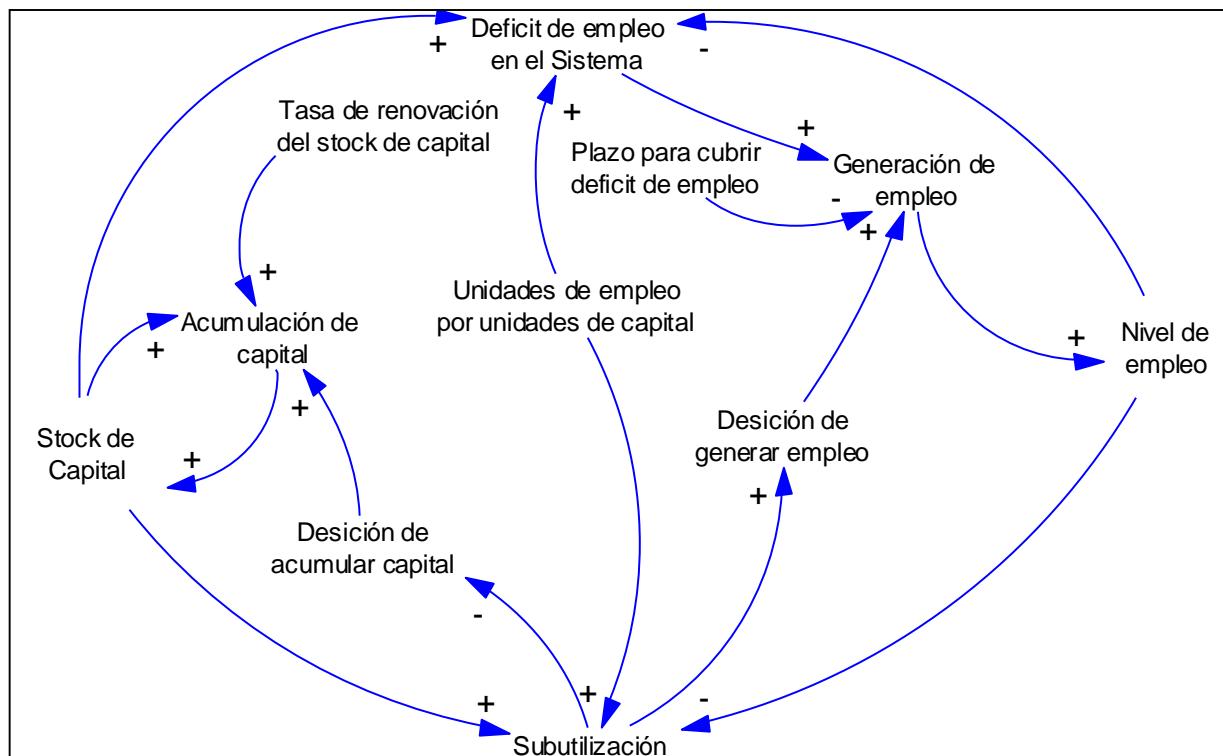
2.



R: La sumatoria de los signos negativos es par, el ciclo es positivo.

5.13 Ejercicios 2

1. Del siguiente esquema determine lo siguiente:
 - ¿Cuántas variables contiene el diagrama?
 - Determine los ciclos cerrados.
 - ¿Cuántas causas-efectos contiene?



Capítulo 6

Diagramas de Flujos Dinámicos (DFD)

Capítulo 6

Diagramas de Flujos Dinámicos (DFD)¹

Los objetivos de este capítulo son:

- Conocer todos los conceptos relacionados a los flujos Dinámicos
- Aplicar estos conceptos en ejemplos simples.

6.1 Introducción

Convertir un diagrama de ciclos causales DCC (también conocido como diagrama causal) en un programa computacional, para poder realizar simulaciones, es una tarea muy difícil ya que el DCC es un modelo conceptual muy genérico del sistema en sí. Para poder realizar simulaciones, se necesita contar con un modelo conceptual que sea más explícito, riguroso y sistemático. Esta técnica de modelado se llama *diagrama de flujos dinámicos* y divide las variables que existen en el sistema en dos tipos: *Niveles* y *flujos*². Los niveles poseen una cantidad que varía a través del tiempo y los flujos poseen tasas o tasas que también varían a través del tiempo. Una estructura básica de niveles y flujos alternados parece representar la naturaleza de los sistemas dinámicos. Los niveles determinan las decisiones que controlan los flujos. Los flujos causan cambios en los niveles. Estos niveles y flujos forman seis redes interconectadas que constituyen la actividad del sistema en estudio. Cinco de estas, representan *material*, *órdenes*, *dinero*, *equipo* y *personal*. La sexta, la red de información, es el tejido conector que interrelaciona las otras cinco. Estas redes poseen flujos a través del tiempo.

Cada diagrama de flujos dinámicos de un sistema dinámico contiene cuatro elementos esenciales (véase la figura 6.1)³:

- Niveles
- Flujos
- Funciones de Decisiones
- Canales de Información.

La figura 6.1 introduce estos cuatro componentes. Una fuente del recurso fluye a través de la línea recta, controlada por la función de decisión (figura rectangular con una llave que se abre o cierra para determinar la tasa que fluye a través de la tubería). Esta tasa cae en un nivel y, este a su vez, cae directamente en un desperdicio. La lectura del nivel

¹ Basado en los conceptos del libro de Jay W. Forrester “Industrial Dynamics”. Diagramas de Forrester.

² Los flujos ocurren a ciertas tasas o tasas.

³ Existen otros que tienen un carácter complementario y se tratan más adelante en este capítulo.

se le envía a la función de decisión para aumentar o disminuir la rata que fluye. Asumamos que lo que fluye es agua.

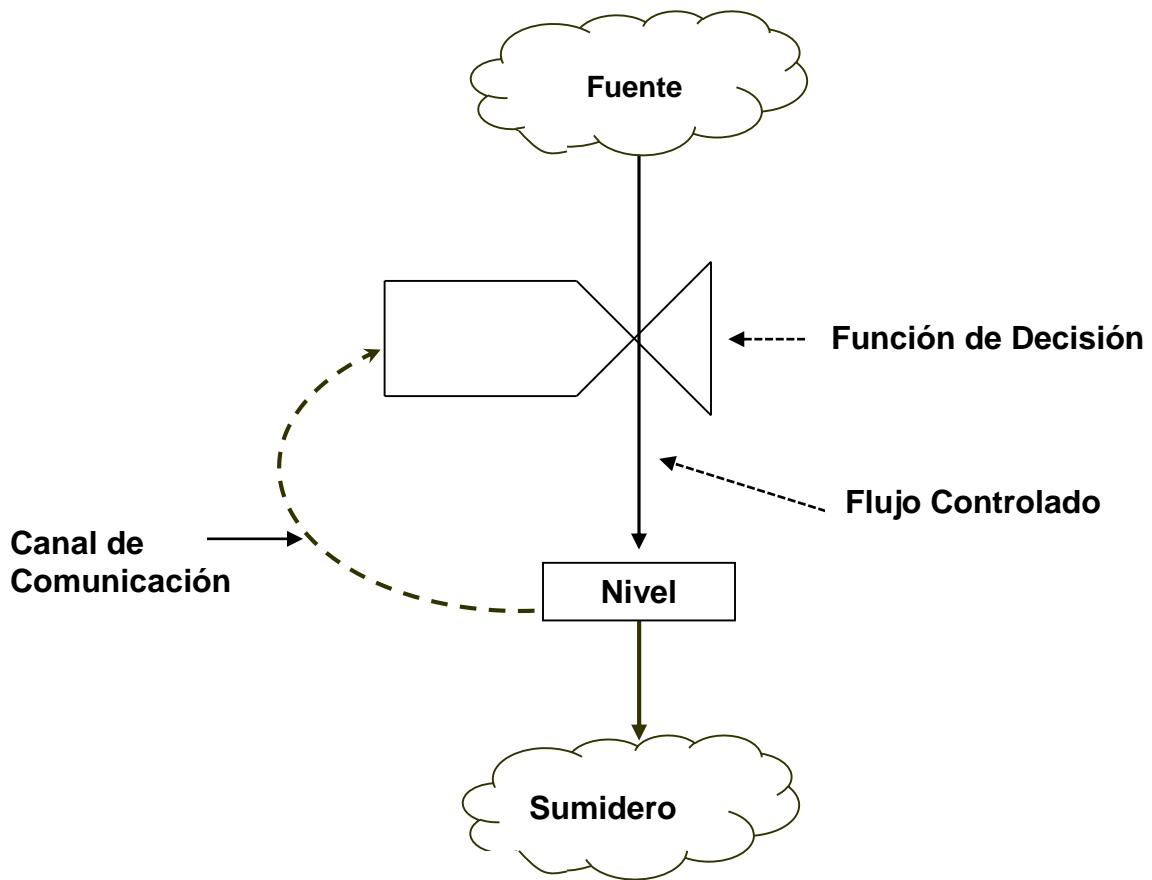


Figura 6.1: Diagrama de Flujo Dinámico o Diagrama de Forrester

La fuente es un lago que posee una tubería que extrae el agua y la envía por medio de una tubería a una llave que determina la rata (cantidad de agua que pasa por minuto por la tubería de 24 pulgadas. Esta agua es utilizada por una ciudad (en este caso, el nivel) y dependiendo de lo que se consume se abre más o menos la llave que aumenta o disminuye la rata de flujo hacia la ciudad. Después de que el agua es consumida en la ciudad, se envía por el alcantarillado hacia el mar (desperdicio)

6.2 Componentes de un Diagrama de Flujos Dinámicos

Una representación gráfica de un sistema de ecuaciones es muy deseable. A continuación, se introducen los componentes de un diagrama de flujos dinámicos DFD. El diagrama se convierte posteriormente en un juego de ecuaciones escritas en un lenguaje como DYNAMO o VENSIM. Un diagrama que muestra las interrelaciones entre las ecuaciones ayuda a dar claridad a la formulación del sistema. Mucha gente visualiza mejor las interrelaciones cuando éstas son mostradas en un diagrama que cuando son hechas de una mera lista de ecuaciones. Un diagrama de flujos dinámicos, debidamente

construido, es mejor que un juego de ecuaciones para comunicar la estructura de un sistema a muchos usuarios y gerentes, esto quiere decir que representa una transición intermedia entre una descripción verbal y un juego de ecuaciones⁴.

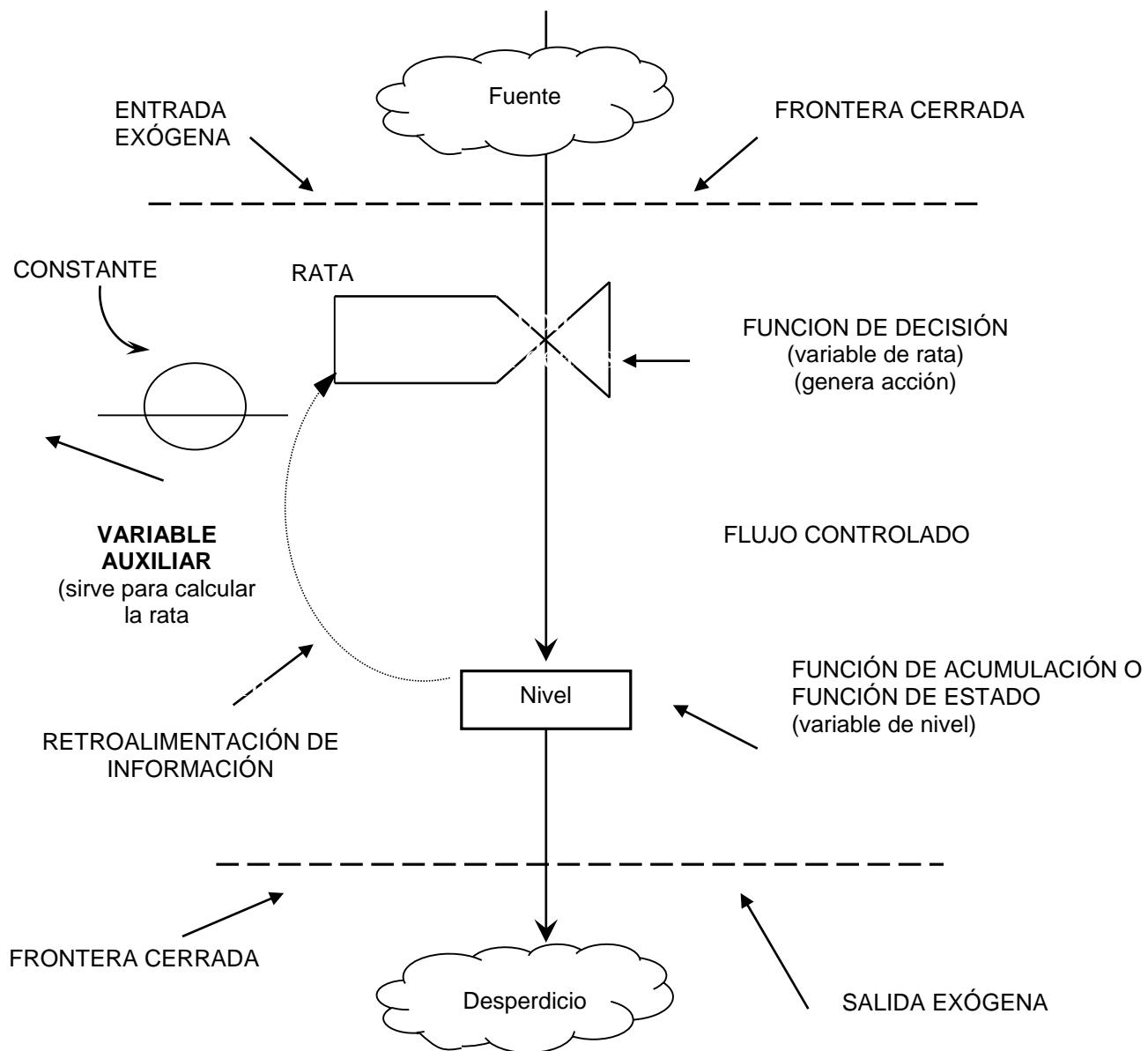


Figura 6.2: Componentes de un DFD

Un diagrama de flujos dinámicos de un sistema interrelacionado se desarrolla primero antes de desarrollar el juego de ecuaciones del sistema. Muchos principiantes intentan desarrollar las ecuaciones primero y luego agregarlas al modelo. Tal negligencia en el

⁴ Forrester Jay, Principles of Systems. Wright-Allen Press. 1976.

diagrama hace perder la claridad, hasta una etapa posterior, ventaja que un diagrama puede proveer en la preparación inicial de la formulación de una ecuación.

La figura 6.2 presenta los componentes más importantes los cuales se explican a continuación. Esta sección describirá un juego de símbolos estándares para los diagramas de flujos de modelos dinámicos. El simbolismo para la presentación por diagramas está basado en escogencias arbitrarias, seleccionadas para recalcar y aclarar aspectos particulares de una situación.

El sistema de símbolos usado aquí muestra la existencia de las interrelaciones en el mismo. Distingue niveles de ratas separando los seis flujos del sistema (información, material, órdenes, dinero, personal y equipo) uno de otro. Revela qué factores intervienen en cada función de decisión (rata), pero el diagrama no revela que las relaciones funcionales están dentro de las funciones de decisión. Para una interacción de naturaleza específica entre los factores que intervienen en una decisión, el diagrama lleva el número de la ecuación, y con esto nos referimos a la ecuación pertinente. El diagrama puede constar de los siguientes componentes:

6.2.1 Nivel

Un nivel está representado por un rectángulo, como en la figura 6.3. En la esquina superior izquierda está un grupo de símbolos (IAR) que denotan una variable particular de nivel. En la esquina inferior derecha está el número de la ecuación para vincular el diagrama a las ecuaciones.

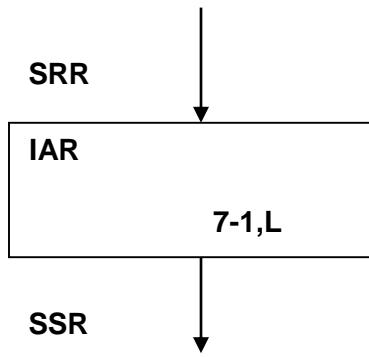


Figura 6.3: Nivel

No se necesita hacer ninguna distinción en el rectángulo para los diferentes flujos del sistema porque las líneas de flujo entrantes y salientes identifican el tipo de flujo (aquí la línea sólida es para el flujo de material). La punta de la flecha muestra la dirección del flujo ya sea hacia, o desde un nivel. Los símbolos al lado de las líneas de flujo identifican las ratas de flujo (a menos que, se muestren en la función de control de decisión).

6.2.2 Rata o Flujo

Un flujo ocurre hacia dentro y hacia fuera de un nivel y posee una rata. El símbolo para una línea de flujo se distingue por los seis tipos de sistemas de flujo, como se muestra en la figura 6.4.

Los tipos de líneas han sido diferenciados para sugerir el tipo de flujo representado y facilitar la lectura del modelo:

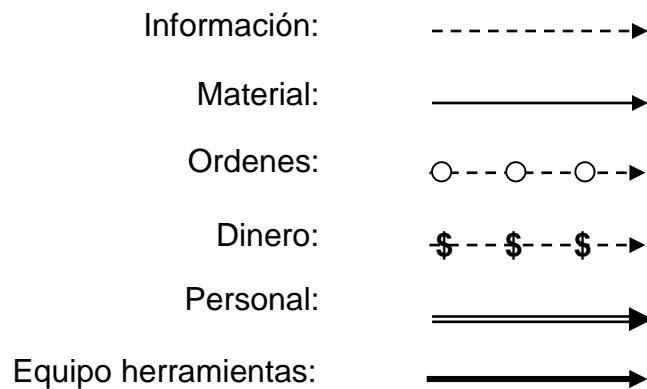


Figura 6.4: Tipos de Flujos

6.2.3 Funciones de Decisión (Ecuaciones de Rata)

Las funciones de decisión determinan la rata del flujo. Ellas actúan como válvulas o llaves en los canales de flujo, como se muestran en la figura 6.5. Dos formas equivalentes son mostradas. El símbolo (SSR) muestra el flujo que está siendo controlado y la información de entrada (solamente información domina en una función de decisión o ecuación de rata) que determina la rata de flujo. El número de la ecuación que define a la rata está dado por (7-5, R).

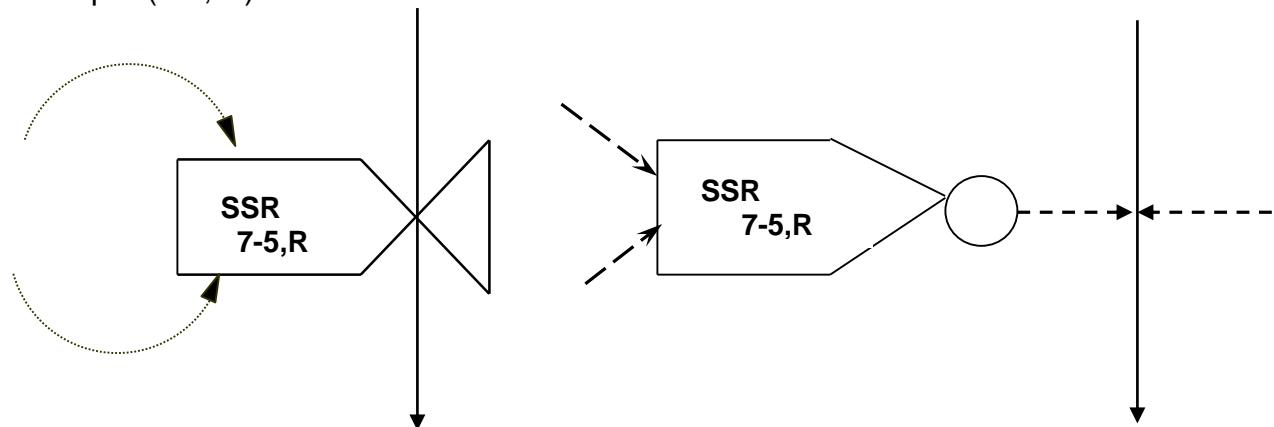


Figura 6.5: Función de Decisión (2 formatos)

6.2.4 Fuentes y Desperdicios

A menudo una rata de un flujo cuya fuente o desperdicio yace fuera de las consideraciones del modelo, debe ser controlada. Por ejemplo, un flujo de órdenes de pedidos de materiales debe empezar en alguna parte, pero la claridad de la terminología del sistema de flujos no permite una mera extensión de las líneas de información a las líneas de los símbolos de órdenes. Las órdenes se consideran apropiadamente cuando empiezan como una provisión de papel en blanco (la fuente) que nosotros no debemos de tratar en el sistema dinámico. Así mismo, las órdenes que han sido llenadas deben ser descartadas del sistema a un archivo de órdenes completadas (desperdicio) que generalmente no tienen características dinámicas significativas.

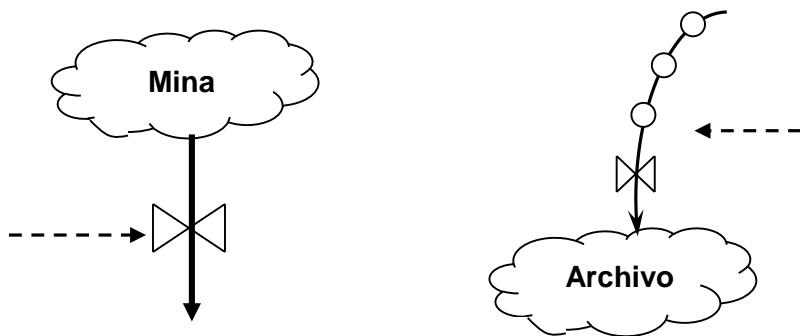


Figura 6.6: Fuente y Desperdicio

En el proceso de manufactura nosotros algunas veces estamos justificados si asumimos que los materiales están disponibles fácilmente y que las características de las fuentes no entran dentro del comportamiento del sistema. En tales circunstancias el flujo controlado es de una fuente o a un desperdicio a la que no se le da ninguna otra consideración en el sistema como se muestra en la fig. 6.6.

6.2.5 Despegue de Información

Los flujos de información conectan muchas variables en el sistema. El despegue de un flujo de información no afecta la variable donde la información es tomada. Tal como en la figura 6.7, un despegue de información es mostrado por un pequeño círculo en la fuente, y en el curso, por una línea de guiones de información.

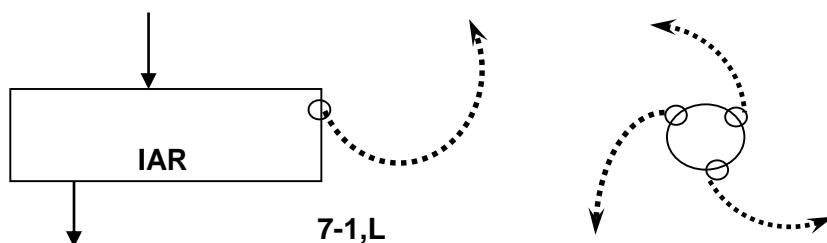


Figura 6.7: El Despegue de Información de un Nivel y una Auxiliar

6.2.6 Variable Auxiliar

Variables auxiliares son conceptos que tienen que ser subdivididos fuera de las funciones de decisión porque ellas tienen un significado independiente. Descansan en los canales de flujo de información entre los niveles y las funciones de decisión que controlan las rutas. Ellas pueden ser algebraicamente sustituidas dentro de las ecuaciones de rata, o sea que no son indispensables, pero simplifican la lectura del diagrama. Las variables auxiliares son mostradas por círculos tal como lo muestra la figura 6.8. Dentro del círculo están el nombre de la variable y el número de la ecuación que la define. Las entradas de la línea de información muestran de donde la variable auxiliar depende (niveles u otras variables auxiliares). La salida de flujos es siempre un despegue de información. La variable auxiliar no es una integración tal como lo es un nivel.

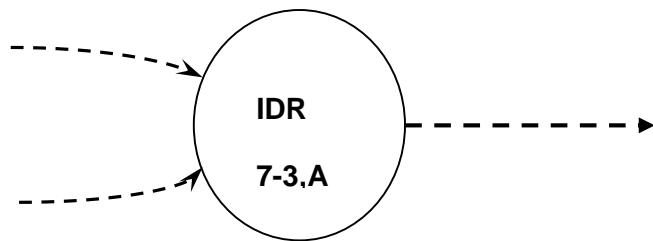


Figura 6.8: Variable Auxiliar

No tiene que guardarse un valor numérico proveniente de un paso computacional de tiempo, al siguiente. Cualquier número de línea de información puede entrar o salir.

6.2.7 Parámetros (Constantes)

Muchos valores numéricos, al describir las características de un sistema, son considerados constantes, al menos durante la duración de una corrida o computación de un modelo simple. Estos son mostrados por una línea arriba o abajo del símbolo de la constante, con despegue de información tal como lo muestra la figura 6.9.



Figura 6.9: Parámetros (Constantes)

6.2.8 Variables en otros Diagramas

Muchos de los diagramas de un sistema están frecuentemente divididos en secciones. La figura 6.10 muestra el manejo de fuentes y la terminación de líneas de flujo que están en otros diagramas. El nombre de la variable y su número de ecuación, se proporcionan; el número de la página del otro diagrama tomado es mostrado en adición.

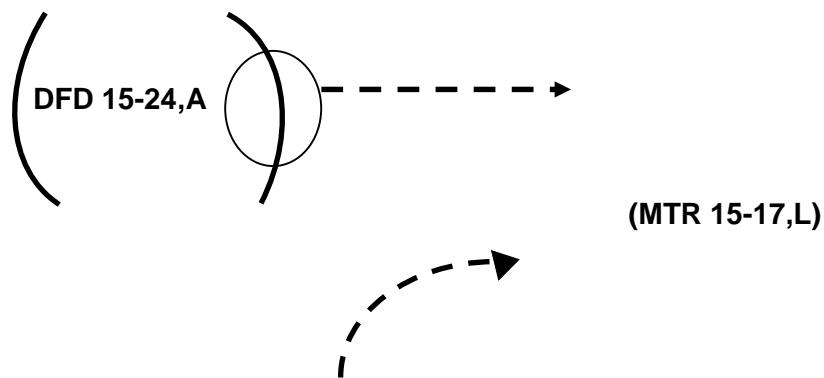


Figura 6.10: Variables que aparecen en otros Diagramas

6.2.9 Retrasos

Los retrasos exponenciales pueden ser representados por una combinación de niveles y ratas de flujo. Sin embargo, los retrasos son encontrados tan frecuentemente, que un símbolo abreviado es necesario, tal como lo muestra la figura 6.11. Este símbolo toma el lugar de los tres niveles con ratas interconectándose (para un retraso exponencial de tercer orden). En la caja, D3 indica un retraso de tercer orden, D1 podría indicar un retraso de primer orden.

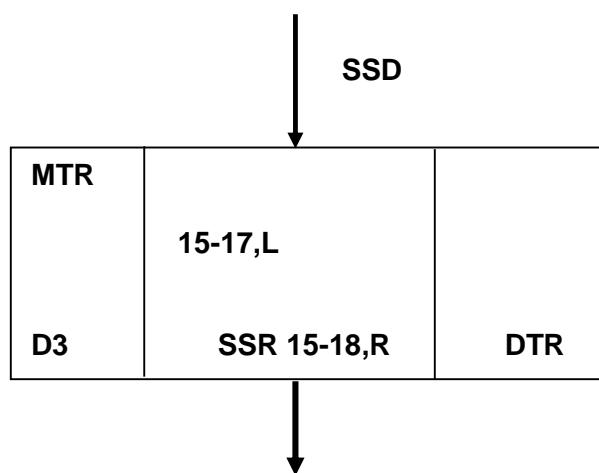


Figura 6.11: Símbolos para Retrasos Exponenciales

SSD	: rata de entrada
MTR	: cantidad (nivel) en tránsito
15-17, L	: ecuación para el nivel en tránsito
D3	: orden de retraso
15-18, R	: ecuación que define a la rata de salida
DTR	: tiempo constante para el retraso
SSR	: rata de salida

6.2.10 Tren de Vagones

Para muchos propósitos, sobre todo situaciones históricas, nosotros queremos extraer información pasada. Para esto, cada una de las líneas de un tren de vagones lineal o cíclico es efectiva. El término "tren de vagones" connota vagones segregados que se mueven con el tiempo. Aquí el movimiento es avanzando el contenido de las cajas al próximo intervalo específico. Los procesos necesarios para complementar las operaciones del tren de vagones son provistos automáticamente por el compilador de DYNAMO.

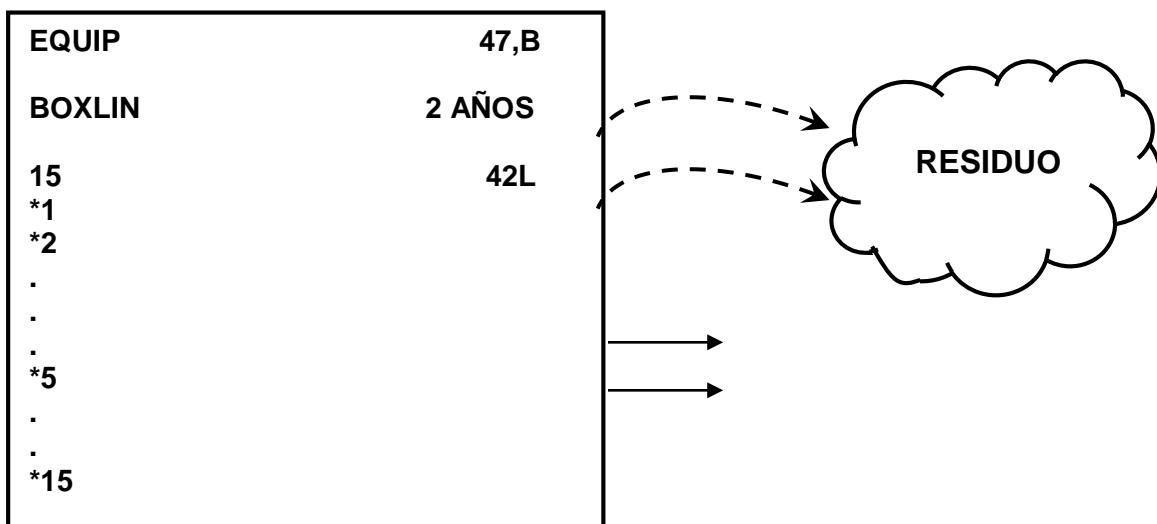


Figura 6.12: Tren de Vagones Lineal

La figura 6.12 muestra un tren de vagones lineal en donde el contenido del vagón es movido hacia abajo un paso en los intervalos específicos y el contenido del último vagón es descartado. Cada vagón tiene una designación que puede ser usada en las ecuaciones de la misma manera como cualquier otra variable. El ejemplo muestra como representar los treinta años de vida de un equipo de capital, seccionado en bloques de dos años. Un tren de vagones de información paralela podría tener productividad separada para indicar cada año del equipo.

BOXLIN : especifica una progresión lineal con lo descartado de la última caja
15 : número de cajas en el tren
2 años : intervalo de avance

El flujo de equipo es mostrado en el primer vagón y es acumulado por la ecuación del nivel 42.

Es descartado automáticamente de la caja.

Los despegues de información son mostrados entre las cajas 1 y 2.

Vagones individuales son identificados por EQUIP * 1, EQUIP * 2, etc.

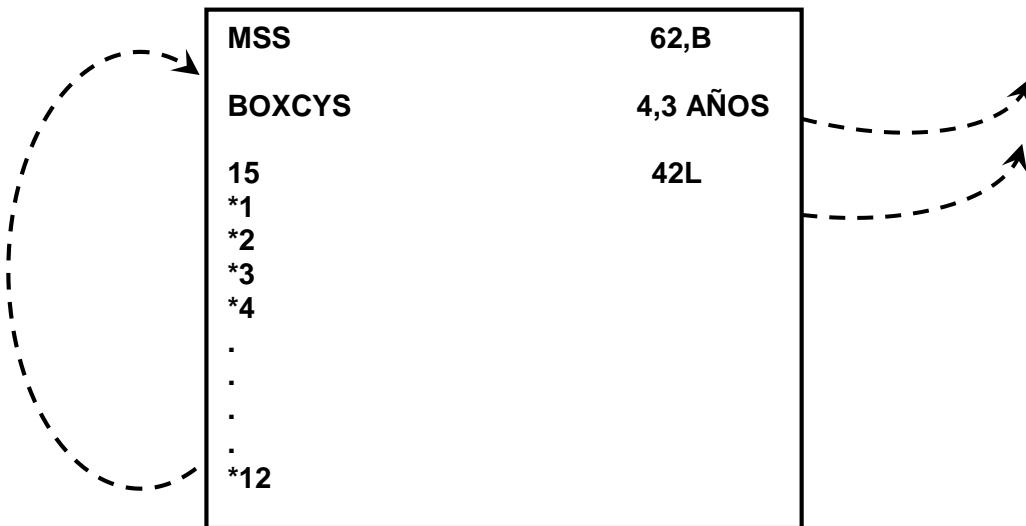


Figura 6.13: Tren de Vagones Cíclico

La figura 6.13 muestra un tren de vagones cíclico en donde el contenido del último vagón es reinsertado en el primer vagón. El ejemplo mostrado puede ser usado para generar el promedio de las ventas mensuales correspondientes a los meses de los años pasados. Tales valores mensuales pueden entrar en las decisiones que son basadas en consideraciones de estación.

MSS : nombre genérico para el tren entero

62,B : número de ecuación que indica en alto las características del tren de vagones

BOXCYS : específica recirculación del último vagón en el primero

- 12 : número de cajones en el ciclo
 4,3 : semanas, intervalos de avance

Los flujos de información van hacia el primer vagón (típico en un cálculo de promedio). Los despegues de información provienen de la caja 2 y 3. Vagones individuales son designados en ecuaciones por MSS * 1, MSS * 2, etc.

6.3 Seis Redes Interconectadas

En esta sección se presentan las seis posibles redes que pueden ser encontradas en un modelo: *órdenes, materiales, dinero, personal y equipo*, todas interconectadas por la red de *información*. Esta particular subdivisión de seis redes puede aparecer en varios segmentos separados: la clase de materiales en un segmento de la red de materiales puede ser diferente y no mezclada, para nuestros propósitos, con otra clase; por lo tanto, los niveles de un segmento no pueden conectarse por rutas de flujos con los niveles de otro segmento.

La red de información puede extenderse de un nivel en cualquiera de las seis redes a una rata de la misma o de cualquiera otra red. La red de información es por lo tanto única y está en una posición superior relativa a las otras cinco. Esto puede ser contrastante con la posición superior que se le asigna frecuentemente al dinero en el análisis económico. Nosotros podemos extender el concepto de dinero más allá de cualquier significado razonable para hacer que éste sirva a los propósitos que se quieran lograr por una red de información distinta.

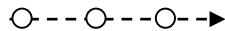
6.3.1 Red de Materiales

Aquí nosotros incluimos todas las rutas de flujo y almacenamiento de bienes físicos, ya sea materiales en bruto, en proceso, inventarios, o en productos terminados.



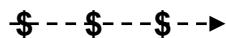
6.3.2 Red de Órdenes

Aquí son órdenes para artículos, requisiciones para nuevos empleados y contratos para nuevos espacios para la planta. Las órdenes son el resultado de decisiones que pueden ser no ejecutadas dentro del flujo de una de las otras redes. En general, ellas forman una cadena entre decisiones abiertas y resultados en la forma de decisiones implícitas en los materiales, dinero, personal o red de capital de equipo. Las órdenes en algunas ocasiones se dificultan para distinguirlas de los conceptos en la red de información. Sin embargo, las órdenes comúnmente tienen una identidad lo suficientemente clara y una posición importante en el sistema que un reconocimiento por separado parece garantizado. Las redes de órdenes pueden usualmente ocurrir en fragmentos que conectan la decisión de generación de órdenes a la acción de relleno de órdenes.



6.3.3 Red de Dinero

El dinero es usado aquí en el sentido de contado. El flujo de dinero es la actual transmisión de pagos entre los niveles de dinero.



Desde el punto de vista del modelo de operaciones de una compañía, el balance del banco es un nivel de dinero. Las cuentas por cobrar no serían dinero, a pesar de que ellas son vendibles como un inventario de bienes; ellas pueden ser llevadas en la red de información como una información que nos dice nuestros derechos al recibo de pago. Los precios también son información y no parte de la red de dinero.

En el punto de relleno de órdenes se crean tres flujos: primero, el flujo de órdenes completas fuera del registro de órdenes no llenas; segundo, el flujo de bienes físicos del proveedor al cliente; y el tercero es el flujo de la información a las cuentas por cobrar el cual es el producto del precio multiplicado por la tasa de flujos de bienes. El nivel de cuentas por pagar al cliente, junto con la factura de pagos diferidos y sus habilidades para pagar, determinan la tasa de flujo de dinero de la cuenta del cliente en el banco al proveedor. Esta tasa de flujo aumenta el nivel de la cuenta del proveedor en el banco, y la información acerca de la tasa baja el nivel de las cuentas por cobrar.

Diferentes puntos de vistas pueden conducir a diferentes tratamientos de información y dinero. La cuenta de banco tratada por la compañía como un nivel de dinero, es para el banco una cuenta por pagar. La cuenta no contiene dinero, pero es una información concerniente a los derechos de la compañía para hacer retiros. La representación de dinero, crédito e información deberá requerir un tratamiento más cuidadoso a medida que los refinamientos del sistema bancario y monetario nacional son incorporados a más modelos dinámicos comprensivos.

6.3.4 Red de Personal

Muchos efectos dinámicos importantes en la compañía surgen de las políticas y comportamiento de patrones circundantes de la adquisición y utilización de personas.



La política de la compañía, contratos de unión obrera y la disponibilidad de personas establecen una estructura que afecta los cambios de niveles de empleos. La medida de tiempo de tales cambios puede interactuar con otras partes de la compañía para producir los resultados inesperados.

En la red de personal tratamos con personas como individuos que se pueden contar, no como hombre-hora de trabajo. El hombre-hora por semana es un producto de los hombres multiplicados por la cantidad de trabajo por semana. Las diferentes consideraciones que controlan los cambios en el número de hombres de aquellos que

circunscriben los cambios son la cantidad de trabajo por semana. En la mayor parte de las situaciones necesitamos distinguir los hombres en la red de personal, de las variables que están en la red de información, tales como la cantidad de trabajo por semana y la productividad por hombre-hora.

6.3.5 Red de Equipos

Incluye espacio de la fábrica, herramientas y el equipo necesario para la producción de bienes o servicios. Esta red describe los caminos y vías en que la fábrica y las maquinarias funcionan, las existencias del capital de equipo disponibles, qué parte (para medir el nivel de productividad) de nuestro capital existente está en uso en cualquier instante, y la rata de desgaste del capital de equipo.



Dentro de un sector productor de bienes de un capital se debe notar que la salida de capital de equipo es un resultado, entre otras cosas del material que fluye dentro de la facilidad de la producción del capital de equipo. El productor del capital de equipo tiene sus propias existencias de equipo, usado para el proceso de producción.

6.3.6 Red de Interconexión de Información

La red de información es una secuencia de ratas y niveles alternados. Esta es ascendida a una posición superior a las otras cinco redes, porque ésta es la interconexión de tejidos entre cada una de ellas. En general la red de información empieza en niveles y ratas de las otras cinco redes y termina en las funciones de decisiones de aquellas cinco redes. Esta transfiere el nivel de información para los puntos de decisión, y la rata de información en las otras cinco redes a los niveles de la red. Los niveles y ratas de flujo existen dentro de la misma red de información. Por ejemplo, la información acerca de la actual rata común de órdenes de entrada es promediada para producir el nivel de promedio de ratas de órdenes de entrada. Este es un nivel en la red de información y sería usualmente una de las entradas para una ordenación de decisión en la red de orden.



El valor de una información variable no debe ser confundido con la verdadera variable que ésta representa. Las órdenes usualmente van dentro de un archivo de órdenes vacías; la información acerca del nivel de rata de órdenes promedio y acerca del nivel de órdenes vacías debe ir a muchos sitios en la compañía. La información, frecuentemente, será retrasada, como los flujos en las otras cinco redes. Como discutimos antes, la información puede contener ruido y confusión. La información no es necesariamente idéntica al valor de la variable que ella presenta.

La red de información contendrá la generación de varios conceptos que son entradas para tomar decisiones, tal como el nivel de inventario deseado, el tamaño proyectado de plantas necesarias, pronóstico de ventas, el nivel deseado de empleados, y el conocimiento de los resultados de las investigaciones. La mayor parte de un modelo

estará usualmente dentro de la red de información. La información es la esencia de la toma de decisiones. Este es el enlace que hace que las otras cinco redes actúen sobre cualquier otra.

En muchos modelos económicos, y en nuestras actitudes hacia la contabilidad, hay frecuentemente un esfuerzo para visualizar dinero en esta central, una relación de enlace para las otras partes de la actividad industrial. La red de dinero no provee entradas adecuadas para la creación de verdaderas decisiones administrativas y económicas. La red de dinero constituye un resumen de las transacciones pasadas y actos como una restricción a futuras decisiones; pero no es guía suficiente para tomar esas decisiones.

La observación cuidadosa de la red de información de una organización es esencial para entender el carácter verdadero de esta red. Su comportamiento dinámico no puede ser representado sin un tratamiento apropiado de los flujos, niveles, retrasos, ruidos y distorsiones que se encuentran en los canales de información que entrelazan los componentes organizacionales separados.

6.4 Estructura de los Sistemas Dinámicos

Todo modelo dinámico que se diseñe, debe cumplir con las siguientes características:

Límites Cerrados

En concepto un sistema de retroalimentación es un sistema cerrado. Su comportamiento dinámico surge dentro de su estructura interna. Cualquier interacción, que es esencial al modo de comportamiento que es investigado, debe ser incluida dentro de los límites del sistema.

Las decisiones siempre están dentro de los ciclos de retroalimentación

Cada decisión es hecha dentro de un ciclo de retroalimentación. La decisión controla la acción que altera los niveles de los sistemas que influyen en las decisiones. Un proceso de decisión puede ser parte de más de un ciclo de retroalimentación.

Ciclo de Retroalimentación - Elemento estructural del sistema

El ciclo de retroalimentación es el elemento estructural básico en los sistemas dinámicos. El comportamiento dinámico se genera por la retroalimentación. Los sistemas más complejos son ensamblajes de ciclos de retroalimentación interactivos.

Niveles y Ratas como Subestructura de los Ciclos

Un ciclo de retroalimentación consiste de dos tipos distintivos y diferentes de variables: los niveles (estados) y las ratas (acciones). Excepto por las constantes, éstas dos son suficientes para representar un ciclo de retroalimentación. Ambas son necesarias.

Los niveles son Integraciones

Los niveles integran (acumulan) los resultados de acciones en un sistema. Las variables de nivel no pueden ser cambiadas instantáneamente. Los niveles crean una continuidad del sistema entre puntos de tiempo.

Los niveles solo son cambiados por las ratas

Una variable de nivel es calculada por el cambio, debido a las variables de rata, que altera el valor previo del nivel. El valor anterior del nivel es transmitido del período previo, este es alterado por las ratas que fluyen a través del intervalo de tiempo que interviene. El valor presente de una variable de nivel puede ser calculado sin los valores presentes o previos de cualquier otra variable de nivel.

Los niveles no se distinguen por unidades de medidas

Las ecuaciones de rata (declaración política) de un sistema son una simple forma algebraica; ellas no involucran tiempo ni los intervalos de solución; ellas no dependen de sus propios valores pasados.

Las unidades de medida de una variable no distinguen entre un nivel y una rata. La identificación debe reconocer la diferencia entre una variable creada por integración y una que es una declaración política en el sistema. Las ratas no se pueden medir instantáneamente

Ningún flujo de rata se puede medir excepto como un promedio en un período de tiempo.

Ninguna rata puede, en principio, controlar otra rata sin la intervención de una variable de nivel.

Las ratas dependen sólo de niveles y constantes

El valor de una variable de rata depende sólo de las constantes y de los valores presentes de las variables de nivel. Ninguna variable de rata depende directamente de otra variable de rata.

Las variables de nivel y las variables de flujo deben alternarse

Cualquier camino a través de una estructura de sistemas encuentra variables de nivel y rata alternadas.

Los niveles describen completamente la condición del sistema

Sólo los valores de las variables de nivel se necesitan para describir por completo la condición de un sistema. Las variables de rata no se necesitan porque ellas pueden ser calculadas de los niveles.

Meta, Observación, Discrepancia y la Subestructura de un sistema-acción

Una política o una ecuación de rata reconoce la meta local hacia la cual el punto de decisión se esfuerza, compara la meta con la aparente condición de un sistema para detectar la discrepancia, y utiliza la discrepancia para guiar la acción. Las jerarquías en una estructura de sistemas pueden resumirse así:

- Límite cerrado
- Estructura de ciclo de retroalimentación
- Estructura de nivel y rata
- Meta, observación, discrepancia y
- Acción como una subestructura dentro de las ratas.

6.5 Sistemas de Ecuaciones

En la sección dos de este capítulo fue ilustrado un arreglo para la estructura básica de un modelo en términos de niveles (reservados) interconectados por rutas de flujos. Para describir esta estructura general, es necesario tener un sistema de ecuaciones adecuado.

El sistema de ecuaciones debe ser adecuado para describir las situaciones, conceptos, interacciones, y proceso de decisiones que constituyen el mundo de la administración y la economía. El modelo debe ser capaz de llevar nuestros conceptos a la realidad en vez de requerir de simplificaciones violentas e inaceptables para traer dichos conceptos dentro de los límites de la estructura del modelo.

Las ecuaciones que se describirán aquí de un sistema básico, van con la elaboración pequeña de la estructura del modelo ya descrito. Las clases de ecuaciones son discutidas en esta sección; pero no así las formas especiales que las ecuaciones de este tipo pueden asumir.

Básicamente, el sistema de ecuaciones consiste de dos tipos de ecuaciones que corresponden al nivel y a los flujos discutidos con anterioridad. En la sección siguiente serán introducidos otros tipos de ecuaciones, las cuales dan conveniencia y claridad a sistemas complejos que no son fundamentales para la estructura del modelo. Antes de hacer esto, hay que añadir que el tiempo fundamental de secuencia de computación puede ser descrito en términos de niveles y flujos:

6.5.1 Secuencia de Computación

Un sistema de ecuaciones se describe en el contexto de ciertas convenciones que muestran cómo las ecuaciones deben ser evaluadas. Tratamos aquí con un sistema de ecuaciones que controla a medida que pasa el tiempo, las interacciones cambiantes de

un juego de variables. La evolución de este sistema implica que la ecuación será calculada periódicamente para producir los nuevos estados sucesivos del sistema.

Para cada punto en el tiempo, debe haber una secuencia particular de computación implicada por el sistema de ecuación. La secuencia usada aquí es presentada en la figura 6.14. El continuo avance de tiempo es roto dentro de pequeños intervalos de igual longitud **DT** (que representa "delta tiempo" o incremento de tiempo). Por definición, este intervalo debe ser lo suficientemente corto, para que consideremos ratas de flujo constantes sobre el intervalo como una aproximación satisfactoria para ratas de variación en el sistema real. Esto quiere decir, que las decisiones hechas al principio del intervalo no pueden ser afectadas por ningún cambio que ocurra durante el intervalo. Al final del intervalo se calculan nuevos valores de niveles y de éstos, nuevas ratas (decisiones) se determinan para el próximo intervalo.

Es claro que, en principio, podemos seleccionar intervalos de tiempo lo suficientemente cortos para que podamos aproximar cualquier curva a segmentos rectos por encima de los intervalos. La figura 6.15 presenta cada línea recta aproximada⁵.

⁵ Forrester Jay, "Principles of Systems. Wright-Allen Press. 1976.

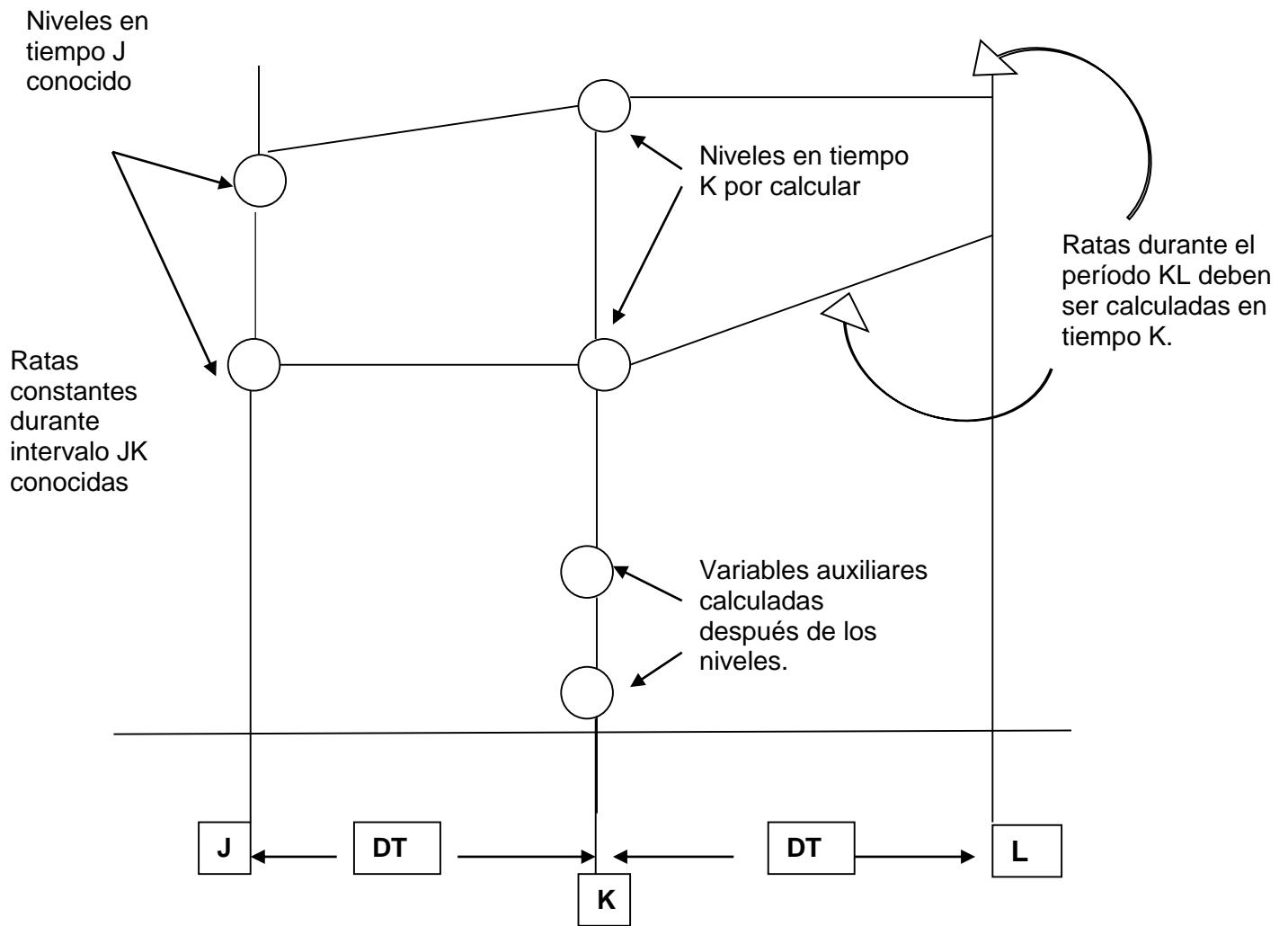


Figura 6.14: Cálculos en Tiempo K.

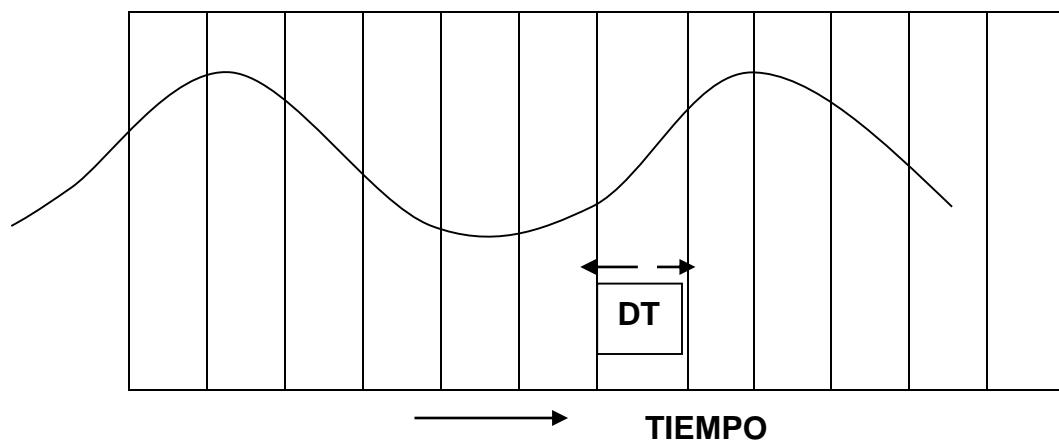


Figura 6.15: Aproximación a una Línea Recta en una Variable de Nivel

Intervalos más cortos y más numerosos pueden producir también una aproximación más cercana a la curva. En la práctica, podríamos lograr intervalos tan pequeños como necesitamos, y que éste tenga aún el tamaño suficiente para retener las sumas de los cálculos hechos dentro de la capacidad de las computadoras de hoy día.

Ahora volvamos a la figura 6.14 en donde puntos sucesivos de tiempo de designación **J**, **K** y **L** son dados. El instante de tiempo **K** se usa para representar el "presente". El intervalo **JK** acabó de pasar, y la información acerca de éste y de los tiempos anteriores, están disponibles para su uso. La información de un tiempo posterior a **K**, semejante al intervalo **KL**, o el tiempo **L**, o más allá no puede estar disponible para su uso en una ecuación que se evalúa al presente tiempo **K**. No existen excepciones para esta indisponibilidad de información futura. Los pronósticos no son información futura; éstos son conceptos presentes acerca del futuro, basado en información disponible del presente y del pasado. Para el propósito de evaluación numérica, la ecuación básica de un modelo está aquí separada dentro de dos grupos: La ecuación de nivel en uno y la ecuación de rata en otro. Para cada paso de tiempo, las ecuaciones de nivel son evaluadas primeramente y los resultados son disponibles para su uso en las ecuaciones de rata (las ecuaciones auxiliares, que se discutirán más tarde, son una conveniencia incidental y son evaluadas entre el nivel y grupos de rata).

Las ecuaciones serán evaluadas en los momentos de tiempo que sean separados por el intervalo de solución **DT**. Las ecuaciones son descritas en términos de pasos generalizados en el tiempo **J**, **K** y **L**, usando la convención arbitraria que **K** representa el punto "presente" en el tiempo en el cual las ecuaciones van a ser evaluadas. En otras palabras, nosotros asumimos que el progreso de la solución será alcanzado justamente en el tiempo **K**, pero que las ecuaciones todavía no tienen solución para los niveles en tiempo **K**, ni para las ratas en el intervalo **KL**.

La ecuación de nivel muestra cómo vamos a obtener los niveles en tiempo **K**, basados en los niveles en tiempo **J**, y en ratas en el intervalo **JK**. En el tiempo **K**, cuando las ecuaciones de nivel son evaluadas, toda la información necesaria está disponible y ha sido llevada adelante desde el punto anterior en el tiempo.

Las ecuaciones de rata son evaluadas en el tiempo presente **K**, después de las ecuaciones de nivel; por eso, tiene disponible los valores de entrada de niveles del presente en **K**. Los valores determinados por las ecuaciones de ratas (decisión) determinan la rata que representa la acción que podremos tomar en el próximo intervalo **KL**. Las ratas constantes implican una rata de cambio constante en los niveles durante un intervalo de tiempo. El declive de las líneas rectas en la figura 6.14 son proporcionales a las ratas y conectadas a los valores de los niveles en los puntos en el tiempo **J**, **K** y **L**.

Después de la evaluación de los niveles en el tiempo **K** y de las ratas en el intervalo **KL**, el tiempo es "indexado". Esto es, las posiciones **J**, **K**, **L** de la figura 6.14 son movidos un intervalo de tiempo a la derecha. Los niveles **K** anteriormente calculados son re etiquetados como niveles **J**. Las ratas **KL** se convierten en ratas **JK**. El tiempo **K**, "el presente", es de este modo avanzado por un intervalo de tiempo de longitud **DT**. La

secuencia entera de computación puede ser repetida para obtener un nuevo estado del sistema en un tiempo que es un **DT** después del estado previo. El modelo traza el curso del sistema a través del tiempo en el medio ambiente (niveles), dirigiendo las decisiones y las acciones (ratas), las cuales, en torno, afectan el medio ambiente. De este modo, las interacciones con el sistema son hechas para seguir la "descripción" que ha sido colocada en las ecuaciones del modelo.

6.5.2 Símbolos en las Ecuaciones

Las variables y las constantes en las ecuaciones son representadas por símbolos (abreviaciones). Consideraciones prácticas hacen deseable el adoptar un estilo estandarizado para los símbolos.

El formato para los símbolos estándares debe permitir cualquier variedad tal que el símbolo lleve una obvia relación a la cantidad que él representa. Por otro lado, la máxima longitud del grupo de letras que se usan como símbolos simples debe ser especificada para facilitar la tarea de programación de la computadora. Las impresoras de las computadoras no ofrecen notación sobrescrita y suscrita, así los símbolos pueden estar compuestos de caracteres en nivel de línea. Para satisfacer las varias demandas, la especificación en un símbolo estándar puede ser usado en este texto:

Para que un símbolo represente una constante o una variable debe consistir de un grupo de seis caracteres o menos, el primero de éstos debe ser alfabético.

En adición a los símbolos básicos, las variables serán seguidas por un punto y el postfijo de tiempo. Los niveles pueden llevar la simple letra **J** o **K** indicando un punto en el tiempo en donde los valores se aplican. Entre los símbolos estándares para niveles podemos incluir los siguientes:

A.J	MEN.J
ABCDEF.K	INV.B.K
B57L.J	B.J
CASH.K	EMPL5.K

Las ratas llevan el postfijo **JK** o **KL** indicando el precedente o sucediente intervalo y puede incluir:

AT.JK	D.JK
BCB.KL	D27.KL
EFF48C.KL	MPM.KL

El compilador **DYNAMO**⁶, para la simulación de la conducta de los modelos, ha sido diseñado para ser consistente con las anteriores convenciones para los símbolos.

⁶ Este lenguaje de simulación fue el primero en esta disciplina. Hoy día utilizamos Stella, Vensim, entre otros.

6.5.3 Notación del Tiempo en las Ecuaciones

Una convención puede ser adoptada para indicar una notación del tiempo tal que podamos especificar el momento en el tiempo donde un valor numérico es aplicado. Frecuentemente en los escritos, el tiempo ha sido indicado por pequeños subíndices bajo la línea de impresión, éstos no son acoplados en una máquina de escribir y no en todas las impresoras de las computadoras. Para designar el tiempo nosotros podremos usar una o dos letras mayúsculas asignándolas a una variable y separadas por un punto. Esto es, en el ejemplo previo el nivel de empleados en el tiempo **J** puede ser **EPLTA.J**, el nivel en el tiempo **K** puede ser escrito como **EPLTA.K**. Note que la simple letra usada en un espacio de tiempo, donde los valores de los niveles son calculados, y son corregidos aquí, en momentos distintos.

Ecuaciones de Nivel:

Una ecuación de nivel representa una reserva para acumular las ratas de flujo que incrementan y decrecen el contenido de la reserva. El nuevo valor de un nivel calculado por adición o sustracción del valor previo, el cambio que puede ocurrir durante la intervención del intervalo de tiempo. Nosotros podemos adoptar el siguiente formato para una ecuación de nivel:

$$L_K = L_J + (DT) (RA_{JK} - RS_{JK}) \quad \text{Ec. 5.1,L}$$

L : Nivel (unidad).

L_K : Nuevo valor del nivel que está siendo calculado en el tiempo **K**.(unidad)

L_J : Valor de nivel para el tiempo pasado **J** (unidades).

DT : Longitud del intervalo de solución entre el tiempo **J** y el tiempo **K** (medida de tiempo).

RA_{JK}: Rata que está siendo sustraída del nivel **L** (unidad / medida de tiempo).

RS_{JK}: El valor de la rata sustraída durante el intervalo de tiempo **JK** (unidad / medida de tiempo).

Cualquier número de ratas, una o más, pueden añadirse o sustraerse de un nivel. Esto es sólo la flexibilidad permisible en la ecuación estándar de nivel. La parte derecha de la ecuación debe contener el valor previo del nivel que está siendo calculado. Debe contener también la solución del intervalo **DT** como un multiplicador del flujo de ratas. La ecuación de nivel es sólo un tipo de ecuación que contiene propiamente la solución del intervalo **DT**.

La solución del intervalo **DT** es un parámetro del proceso computado, no un parámetro del sistema real que representa el modelo. Las ratas de flujo del sistema, medidas en "unidad / medida de tiempo" (por ejemplo, libras / horas, u hombres / meses, o calorías / segundo) son acumuladas en pasos o en grupos sobre intervalos de tiempo sucesivos de longitud **DT**. La solución del intervalo **DT**, medido en unidad de tiempo, convierte el flujo de ratas a una cantidad del objeto que fluye. Esto es el producto del flujo de rata multiplicado por el tiempo que crea la unidad correcta de medidas por adición al valor del

nivel. El intervalo de solución puede ser arbitrariamente cambiado (sino se vuelve demasiado largo) sin afectar la validez del modelo. Todas las otras ecuaciones están formulados en términos de la unidad básica de tiempo usada en el sistema real. El intervalo de solución **DT**, no podría aparecer en cualquier otra ecuación que la ecuación de nivel.

Principio: Intervalo de Solución DT en todas las Ecuaciones de Nivel y no Otras.

El intervalo de solución multiplica las ratas y es esencial en la función de acumulación (o integración) de la ecuación de nivel. Todas las otras ecuaciones podrían estar formuladas en términos de unidades de medición de tiempo habitual, en el sistema real. Despues de la formulación del modelo, el intervalo de solución será seleccionado para garantizar estabilidad dentro de los procesos de computación del mismo (para ser distinguido de sistemas de estabilidad).

La ecuación de nivel realiza el proceso de integración. En la notación de cálculos y ecuaciones diferenciales, la ecuación 5.1 podría ser escrita como sigue:

$$L = L_0 + (RA - RS)dt \quad \text{Ec 5.2}$$

L : El valor del nivel en cualquier tiempo t (unidades)
 L_0 : El valor inicial del nivel a $t=0$

el operador indicador de integración o acumulación desde tiempo igual cero hasta tiempo igual t de la diferencia en rata de flujo ($RA - RS$).

RA : El flujo de rata siendo añadida

RS : El flujo de rata siendo sustraída.

dt : El operador diferencial representando la pequeña diferencia infinitesimal en tiempo que multiplica las ratas de flujo. (corresponde a la longitud de los pasos representados por DT).

La ecuación 5.1 es también conocida como una ecuación diferencial de primer orden en las ramas de las matemáticas que trata las integraciones paso a paso.

Ecuaciones de Ratas:

Ecuaciones de ratas especifican cómo los flujos dentro de un sistema son controlados. La entrada a una ecuación de rata son niveles de sistemas y constantes. Las salidas de una ecuación de rata controlan un flujo "A", de o entre niveles. Siguiendo la notación del tiempo en la sección 3, la ecuación de rata es calculada al tiempo **K**, y usando la información de niveles al tiempo **K**, para encontrar el próximo flujo de ratas para el intervalo **KL**. La forma de una ecuación de rata es:

$$R.KL = (\text{Niveles y Constantes}) \quad \text{Ec. 5.3}$$

donde el lado derecho implica cualquier función o relación, de niveles y constantes que describen la regla que controla la rata. Usando las convenciones que tenemos ahora establecidas, algunas de las ecuaciones de rata pueden volverse a escribir para seguir la notación propia para ratas y niveles.

Ecuación $OR = (1/AT) * (DI - I)$ se convierte:

$$OR.KL = (1/AT) * (DI - I.K) \quad \text{Ec. 5.4, R}$$

Ecuación $RR = GO/DO$ se convierte en:

$$RR.KL = GO.K/DO \quad \text{Ec. 5.5, R}$$

Ecuación $SHR = (1/SDT)*S$ se convierte en:

$$SHR.KL = (1/SDT) * (S.K) \quad \text{Ec. 5.6, R}$$

Diferentes a las ecuaciones de nivel, las ecuaciones de rata no están restringidas en su forma excepto por tres prohibiciones ya implícitas en secciones anteriores:

Una ecuación de rata no podría contener el intervalo de solución **DT**. Excepto en las ecuaciones de nivel, el intervalo de solución no tiene significado en la formulación de ecuaciones del modelo. El intervalo de solución surge del proceso de cálculo paso a paso y es la única cantidad que aparece en las ecuaciones que no tienen significado en el sistema real que el modelo representa.

En cuanto a eso, no podría ser una variable al lado derecho de la ecuación, solamente niveles y constantes.

El lado izquierdo de la ecuación contiene una variable de rata siendo definida por la ecuación. El valor de la rata para el intervalo **KL**, está siendo calculado inmediatamente en el término **K**.

Las ecuaciones de ratas son reglas, que dicen cómo las "decisiones" son hechas. La regla (ecuación de rata) es la oración general de cómo la información pertinente es convertida en una decisión (o flujo, o corriente de acción presente - siendo todos términos sinónimos). Las ecuaciones de rata nos dicen cómo el sistema se controla a sí mismo. Las palabras, "reglas" y "decisiones", tienen significados amplios aquí que en su uso común. Ellas van más lejos que las usuales decisiones humanas e incluyen el control de los procesos que están implícitos en la estructura del sistema, y en costumbre y tradición. Una ecuación de rata (o reglamento) puede describir como el flujo en una tubería depende del valor de la posición y la presión del fluido dentro de una organización; o una ecuación de rata puede representar las reglas explícitas que controlan el flujo de información en el sistema real, donde los procesos están automáticamente bajo el control de un programa de computadora digital. Las ecuaciones de ratas son más útiles que las ecuaciones de nivel. Las ecuaciones de ratas muestran la percepción de cómo las

decisiones del sistema real responden a las circunstancias que rodean el punto de decisión.

Ecuaciones Auxiliares:

Muy a menudo, la claridad y significado de la ecuación de rata puede ser mejorado dividiéndola en partes que son escritas como ecuaciones separadas. Esas partes podemos llamarlas ecuaciones auxiliares. La presencia de ecuaciones auxiliares en un modelo, de ninguna manera contradice el concepto de que la estructura del sistema está compuesta sólo de niveles y ratas. Las ecuaciones auxiliares son meramente subdivisiones algebraicas de las ratas. Supóngase que el inventario deseado en una ecuación de ordenamiento de inventario es una variable que depende de la rata de ventas promedio. La ecuación de rata para el ordenamiento y la respectiva ecuación auxiliar para el inventario deseado podría ser:

$$OR.KL = (1/AT)(DI.K - I.K) \quad \text{Ec. 5.7, R}$$

$$DI.K = (WID)(ASR.K) \quad \text{Ec. 5.8, A}$$

OR : rata de orden (unidades / semana)

AT : tiempo de ajuste

DI : inventario deseado

I : inventario

WID: semanas de inventario deseado (semanas)

ASR: rata de ventas promedio (unidades / semana)

En la ecuación 5.8, **WID** es una constante, cuyo valor expresa el inventario deseado en términos de semanas de ventas promedio. La rata de venta promedio, **ASR**, es un nivel. La ecuación 5.8 puede ser sustituida en la ecuación 5.7 para crear la siguiente ecuación de rata que depende sólo de niveles y constantes:

$$OR.KL = (1/AT) [(WID)(ASR.K) - I.K] \quad \text{Ec. 5.9, R}$$

De esta manera la ecuación auxiliar ha desaparecido. Las ecuaciones auxiliares deben ser evaluadas luego de las ecuaciones de nivel, de las cuales dependen, y antes de las ecuaciones de ratas de las que son parte. Cuando una ecuación auxiliar existe, y generalmente serán muy numerosas, la computación se realiza en esta secuencia:

- niveles
- auxiliares
- ratas

A diferencia de las de niveles y ratas, las ecuaciones auxiliares pueden depender de otras ecuaciones auxiliares en una cadena; así que algunos grupos de ecuaciones tendrán que ser evaluadas en un orden particular. Considere las siguientes ecuaciones:

SH.KL = (1/SAT)(IS.K - S.K)	Ec. 5.10, R
IS.K = B.K/SS	Ec. 5.11, A
B.K = (OB.K)(RS)	Ec. 5.12, A
OB.K = (S.K)(SE.K)	Ec. 5.13, A
SE.K = TABLE(TSE, DDR.K, 0, 6, .5)	Ec. 5.14, A

En la secuencia de ecuaciones anteriores, es evidente que la primera es una ecuación de rata debido a la notación de tiempo **KL**. Las otras son ecuaciones auxiliares que tienen el mismo designador de tiempo **K** como las ecuaciones de nivel; pero no son de nivel como aparentan ser debido a que el formato de la ecuación no es el de la ecuación de nivel.

La ecuación 5.14 no es una ecuación algebraica, es más bien una instrucción operacional que expresa que la efectividad de ventas (**SE**) es una función de **DDR**. El formato es aquel utilizado por el compilador DYNAMO para una función de tabla. Los números en la última parte de la expresión nos dicen que la tabla mide los valores del 0 al 6 con un espaciado de 0.5 en la escala de medida de **DDR**.

Empezando desde la última del grupo de ecuaciones anterior, es posible sustituir los valores y a su vez ir reemplazando en las ecuaciones que lo preceden. Empezando con los valores de los niveles de **DDR** y **S**, las ecuaciones auxiliares en el grupo anterior deben ser evaluadas en la secuencia **SE, OB, B e IS**.

Cuando existan cadenas interconectadas de ecuaciones auxiliares, ellas deben ser evaluadas en una secuencia que permita sustituciones sucesivas. Tal secuencia siempre existirá en un sistema formulado apropiadamente. Si existe un ciclo circular de ecuaciones auxiliares esto implica un conjunto de ecuaciones algebraicas simultáneas, una solución en secuencia no existe, y las auxiliares que no son simples enlaces de las ecuaciones de nivel serán encontradas y rechazadas por el compilador DYNAMO como un error de formulación del sistema.

6.5.4 Constantes y Ecuaciones de Valores Iniciales

Constantes: Una constante es representada por un nombre simbólico, le es asignado un valor numérico en una ecuación de constante. Las ecuaciones de constante tienen el designador de **clase C** después del número de la ecuación. Una constante no tiene una variable de tiempo porque no cambia a través del tiempo.

XY.K = (AB)(Z.K)	Ec. 5.15, A
AB = 15	Ec. 5.16, C

El valor de la constante AB es dado en la ecuación de constante 5.15. El número de la ecuación de una constante será dado como una subdivisión en decimal del número de la ecuación primaria en donde la constante aparece por primera vez.

Constantes Calculadas Inicialmente: Es frecuentemente conveniente especificar una constante en términos de otra constante cuando la primera depende de la última y cuando la primera debe cambiar en una corrida de la simulación en donde a la última le es dado un nuevo valor. Supongamos que la constante **CD** siempre será 14 veces el valor de **AB**. **AB** es dada arriba en la Ec. 5.16 donde su valor puede ser alterado del original 15. **CD** se podrá escribir como:

$$CD = (14)(AB) \quad \text{Ec. 5.17, N}$$

El designador de **clase N** (el mismo que para las ecuaciones de valor inicial que dan el valor inicial de niveles) indica que esta ecuación necesita ser evaluada solamente una vez al comienzo de la computación de la simulación, porque las constantes, por su naturaleza y definición, son valores que no deben variar durante ninguna corrida de simulación. El número de la ecuación será una subdivisión decimal de la ecuación donde inicialmente la constante calculada por primera vez, apareció.

Ecuaciones de Valores Iniciales: A todas las ecuaciones de nivel se le deben asignar valores iniciales al comienzo de la computación de simulación. Estas variables de nivel representan la condición completa del sistema necesaria para determinar las rutas de flujo venideras. Toda la historia del sistema que influye en acciones presentes es representada por los valores presentes de las variables de nivel apropiadas en solamente la forma de la historia del sistema existente en el presente, puede ser efectiva. Las recolecciones del pasado son atenuadas y modificadas por el tiempo. Es la presente versión de la historia representada por los valores presentes de los niveles del sistema la que gobierna la acción. Los valores iniciales para las variables de rata no deben y no necesitan ser dados porque ellos son completamente determinados por los valores iniciales de las variables de nivel.

De los valores iniciales de las variables de nivel, las rutas de flujo que inmediatamente siguen al tiempo cero, pueden ser calculadas y con estos valores iniciales y las rutas, los nuevos valores de los niveles al final del primer paso de tiempo pueden ser calculadas. Las ecuaciones de valor inicial llevarán el designador de **clase N** después del número de la ecuación. No se utilizan variables de tiempo. El lado derecho de la ecuación de valor inicial es escrito en términos de valores numéricos, constantes indicadas simbólicamente y los valores iniciales de otros niveles. Dos o más valores iniciales no dependen mutuamente en uno y otro porque resultaría indeterminado. En cuanto a eso, debe ser un punto de partida inicial cualquier serie de valores iniciales que es expresado en términos constantes. Una ecuación de inicialización es habitualmente escrita inmediatamente siguiendo la correspondiente ecuación de nivel:

$$PT.K = PT.J + (DT)(M.JK - N.JK) \quad \text{Ec. 5.18, L}$$

$$PT = 8 \quad \text{Ec. 5.19, N}$$

El número de la ecuación de inicialización estará dado como una subdivisión decimal del número de su ecuación de nivel (como en Ec. 5.18 y Ec. 5.19 arriba). La ecuación de iniciación de arriba también pudo haber sido escrita en términos de constantes como:

$$\begin{array}{ll} PT = (3)(CD) & \text{Ec. 5.20, N} \\ PT = AB & \text{Ec. 5.21, N} \end{array}$$

Esto es permisible y no es ambiguo para expresar un valor inicial de una ecuación de nivel en términos del valor inicial, tan largo como el posterior e independiente del primero. Por ejemplo, el siguiente valor inicial depende del valor inicial en la Ec. 5.19:

$$\begin{array}{ll} RS.K = RS.J + (DT)(ML.JK - NL.JK) & \text{Ec. 5.22, L} \\ RS = PT & \text{Ec. 5.23, N} \end{array}$$

6.6 Resumen

En este capítulo se ha presentado un modelo conceptual permite diagramar con mucho detalle el comportamiento dinámico de un sistema. Para poder hacer esto, fue necesario identificar todas las variables pertinentes y clasificarlas en niveles y flujos. Esta clasificación es trascendental en el mundo dinámico de los sistemas. Es común, que el modelador descubra primordialmente las variables tipo nivel; pero también es muy importante descubrir todas variables de flujo. No obstante, dos variables tipo nivel necesitan tener entre ellas una variable tipo flujo y así se completan las variables faltantes. Normalmente no se consigue un diagrama de flujos dinámico perfecto después del primer intento. Tal como se manifestó en el capítulo 2, será necesario refinar el DFD por medio de validaciones repetidas que certifiquen la calidad y similitud del modelo con el sistema real bajo estudio.

Capítulo 9

Introducción a los Diagramas de Bloques

Capítulo 9

Introducción a los Diagramas de Bloques¹

9.1 Introducción

Existen diferentes formas de representar un *sistema de control*². En el capítulo 5 se presentó el método de diagramas de ciclos causales. En este tendremos la oportunidad de tratar otro que es muy utilizado sobre todo para analizar y diseñar sistemas físicos.

Los diagramas de bloques que se describen a continuación, son una simbología especial para representar sistemas complejos, y es muy utilizado en diferentes ingenierías. La combinación de diagramas de bloques y funciones de transferencia de un sistema físico provee una representación gráfica de la relación causa-efecto entre la salida y la entrada de un sistema.

9.2 Definición de Diagrama de Bloques

Un diagrama de bloques es una representación gráfica simplificada de las relaciones causa y efecto que existen entre las entradas y salidas de un sistema físico. El diagrama suministra un método útil y conveniente para caracterizar las relaciones funcionales entre los diferentes componentes de un sistema de control.

Los diagramas de bloques también pueden ser definidos como herramientas que sirven como auxiliares para la descripción de sistemas tanto en etapas preliminares del análisis como en etapas posteriores.

Para representar la transformación de una variable a otra se utilizan los bloques o cajas negras, los cuales tienen una entrada y una salida. Sobre los bloques correspondientes, se colocan generalmente las funciones de transferencia de los componentes; los bloques están conectados por flechas para indicar la dirección del flujo de señales. Debe notarse que la magnitud de la señal de salida del bloque es la de la señal de entrada, multiplicada por la magnitud de la función de transferencia en el bloque.

Las ventajas de la representación de un sistema por medio de un diagrama de bloques consisten en que es fácil formar el diagrama global de todo el sistema, colocando simplemente los bloques de sus componentes de acuerdo con el flujo de señales, y en

¹ Material revisado gentilmente por el Profesor Lino Ruiz, UTP.

² Ecuaciones Matemáticas y Diagramas de Flujo de Señales.

que es posible evaluar la contribución de cada componente al comportamiento general de todo el sistema.

En general, el funcionamiento de un sistema se puede ver más fácilmente, examinando el diagrama de bloques, que analizando el sistema físico en sí. Un diagrama de bloques contiene información respecto al comportamiento dinámico, pero no contiene ninguna información acerca de la constitución física del sistema. En consecuencia, muchos sistemas disímiles, sin relación alguna entre sí, pueden estar representados por el mismo diagrama de bloques.

A continuación, la figura 9.1 de un elemento de diagrama de bloques. La flecha que apunta hacia el bloque representa la entrada³, y la que se aleja del bloque representa la salida⁴. Estas flechas, comúnmente tienen el nombre de señales.



Figura 9.1: Diagrama de Bloque

La relación de transferencia tiene la siguiente fórmula:

$$\frac{R(s)}{E(s)} = F(s)$$

Un error grave que se podría dar es la mala identificación de las entradas y salidas del sistema. Para cada componente y todo el sistema en general, se debe tener en cuenta que las entradas son las causas de lo que ocurre en un subsistema y las salidas, el resultado del subsistema. De esta forma se llega al concepto de *sistema causal*. También se debe aclarar que se pueden dibujar diferentes diagramas de bloques del mismo sistema, dependiendo sólo del punto de vista del análisis.

A pesar de que todos los sistemas (con una entrada y una salida) pueden ser denotados por un bloque único conectado entre la entrada y la salida, la ventaja del concepto de diagrama de bloques reside en el hecho de que los sistemas de control están compuestos de numerosos elementos no interactivos cuyas funciones de transferencia son determinadas independientemente.

³ Actuación, estímulo o señal de entrada.

⁴ Respuesta o variable controlada.

Un sistema entero puede ser representado por la interconexión de los bloques de elementos individuales, de manera que sus contribuciones al comportamiento total del sistema pueden ser evaluados.

9.3 Simbología Utilizada

Los diagramas de bloques se componen de los siguientes símbolos:

Rectángulo:



Contienen las palabras descriptivas o de identificación de la función de transferencia. Representan los componentes, elementos y subsistemas del sistema.

Flecha:



Describen el flujo de transacciones entre los elementos (rectángulos). La flecha que apunta hacia el bloque indica la entrada y la que se aleja del bloque representa la salida. Normalmente se les llaman *señales*.

Punto de Suma:



Un circuito con una cruz constituye el símbolo que indica si la señal a de sumarse o restarse. Es importante que las cantidades a sumar y restar tengan las mismas dimensiones y las mismas unidades. También se le conoce como *recolector*.

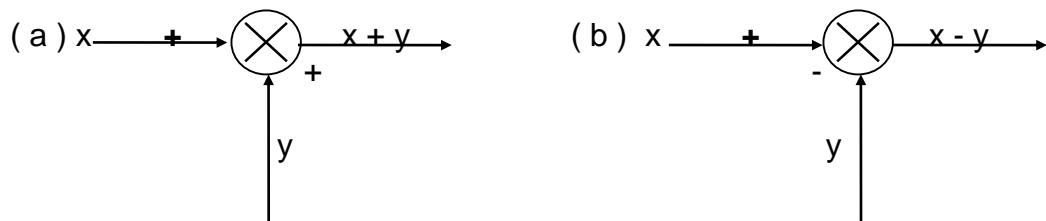
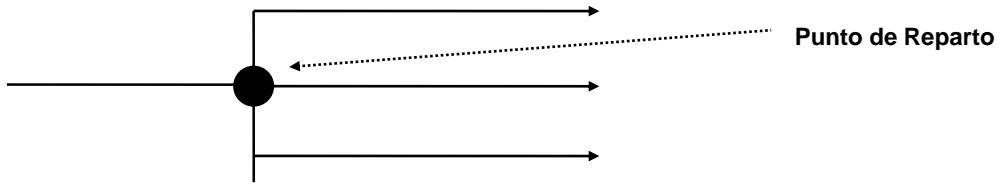


Figura 9.2: Ejemplos de Puntos de Suma

Punto de Reparto:

Un punto negro representa un punto donde una señal se bifurca y su valor se distribuye igualmente por cada bifurcación. También se le conoce como punto de bifurcación.

**Figura 9.3: Punto de Reparto**

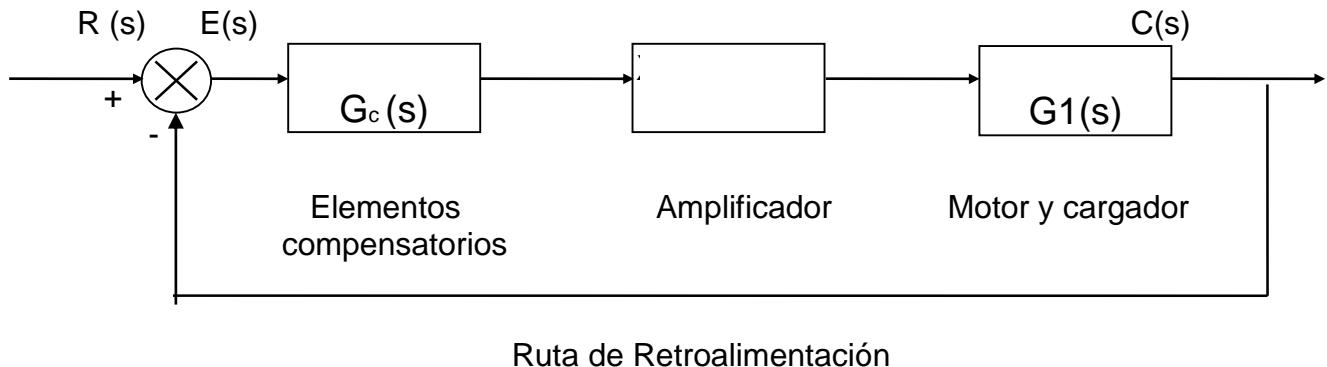
En general, los diagramas de bloques constan de una configuración específica de cuatro tipos de elementos o bloques: rectángulos, puntos de suma, puntos de bifurcación y flechas que representan el flujo unidireccional de señales.

9.4 Tipos de Diagramas de Bloques

9.4.1 Diagrama de Bloques de un Sistema de Control de Retroalimentación

Un sistema entero puede ser representado por la interconexión de los bloques de elementos individuales, de manera que sus contribuciones al comportamiento total del sistema pueden ser evaluados.

El diagrama en bloques de un sistema de control de retroalimentación típico es mostrado en la figura 9.4. Cada bloque representa un elemento independiente del sistema:

**Figura 9.4: Diagrama de Bloques de un Sistema de Retroalimentación**

9.4.2 Diagramas de Bloques de un Sistema de Control con Retroalimentación Multivariable

Un diagrama de bloques de un sistema multivariable puede ser dibujado como se muestra en la Figura 9.5. Este diagrama de bloques solo puede especificar las relaciones de transferencia entre las entradas y las salidas.

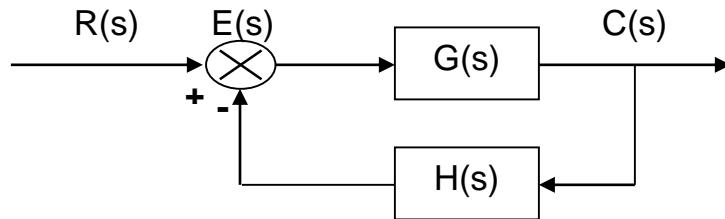


Figura 9.5: Diagrama de Bloques con Multivariables

9.4.3 Relaciones Básicas del Diagrama de Bloques

La mayoría de los sistemas se encuentran formados por diferentes "s" interconectados entre sí. Por lo tanto, es necesario conocer cómo se realizan las interconexiones, además, estas conexiones de las "s" sólo son posibles si las señales son compatibles.

Tipos de Conexiones

9.4.3.1 Conexión en Serie

Este tipo de conexión sólo es posible si las señales son compatibles porque la salida de un bloque pasa a ser la entrada del otro. El diagrama sería el siguiente:

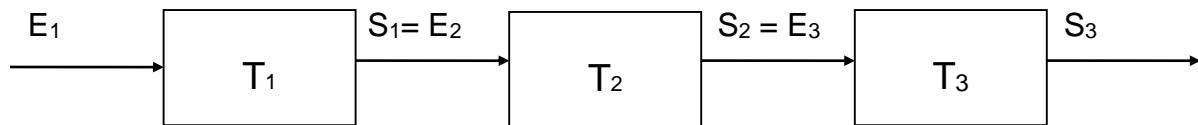


Figura 9.6: Diagrama de Bloques en Serie

Las ecuaciones serían de la siguiente forma:

$$T_1(E_1) = S_1 = E_2$$

$$T_2(E_2) = S_2 = E_3$$

$$T_3(E_3) = S_3$$

$$S_3 = T_1 * T_2 * T_3$$

$$E_1$$

9.4.3.2 Conexión en Paralelo

Utilizando las propiedades del punto de suma puede existir también la conexión en paralelo. Esta se da cuando la entrada es igual para todos los subsistemas existentes. El diagrama para este tipo de conexión es el siguiente:

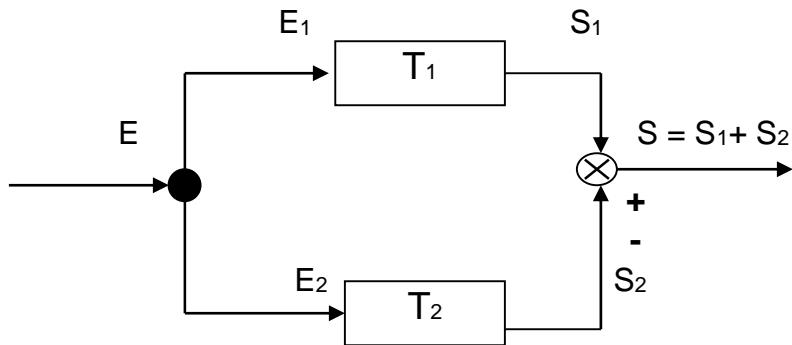


Figura 9.7: Diagrama de Bloques en Paralelo

Las ecuaciones son las siguientes:

$$S = S_1 + S_2$$

$$S_1 = E_1 T_1$$

$$S_2 = E_2 T_2$$

$$S = T_1 E_1 + T_2 E_2$$

$$\frac{S}{E} = T_1 \pm T_2$$

De lo anterior se deduce que varios subsistemas conectados en paralelo producen en las variables de entrada una transformación a cada uno de los subsistemas.

9.5 Funciones de Transferencia

9.5.1 Ciclo Abierto

La relación de la señal retroalimentada $B(s)$ con respecto a la señal de error $E(s)$ se llama función de transferencia de ciclo abierto. Es decir, función de transferencia de trayectoria abierta $= B(s) / E(s) = G(s)H(s)$.

9.5.2 Ciclo Cerrado

La figura 9.8 muestra un ejemplo de diagrama de bloque de un sistema de ciclo cerrado. La salida $C(s)$ es alimentada nuevamente al punto de suma, donde se le compara con la entrada de referencia $R(s)$. La naturaleza de ciclo cerrado del sistema queda claramente indicada por la figura 9.8. La salida del bloque $C(s)$ es obtenida en este caso multiplicando la función de transferencia $G(s)$ por la entrada al bloque $E(s)$. Cualquier sistema de control lineal puede, ser representado por un diagrama de bloques consistente en bloques, puntos de suma y puntos de bifurcación.

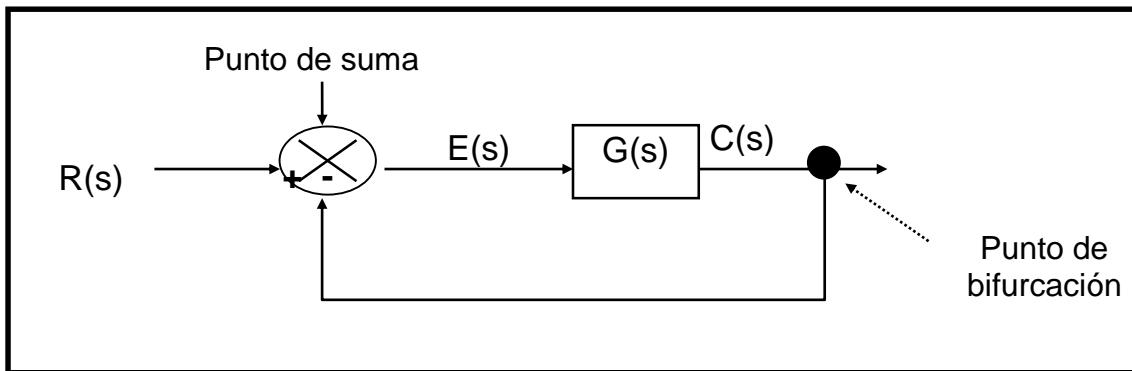


Figura 9.8: Diagrama de Bloque de un Sistema de Ciclo Cerrado

Al inyectar nuevamente la salida al punto de suma para comparación con la entrada, es necesario convertir la forma de la señal de salida a la forma de la señal de entrada. La función de transferencia que relaciona $C(s)$ y $R(s)$ se denomina función de transferencia de ciclo cerrado que se muestra en la Figura 9.9. Esta función de transferencia relaciona la dinámica del sistema de ciclo cerrado a la dinámica de los elementos de alimentación directa y a los elementos de retroalimentación.

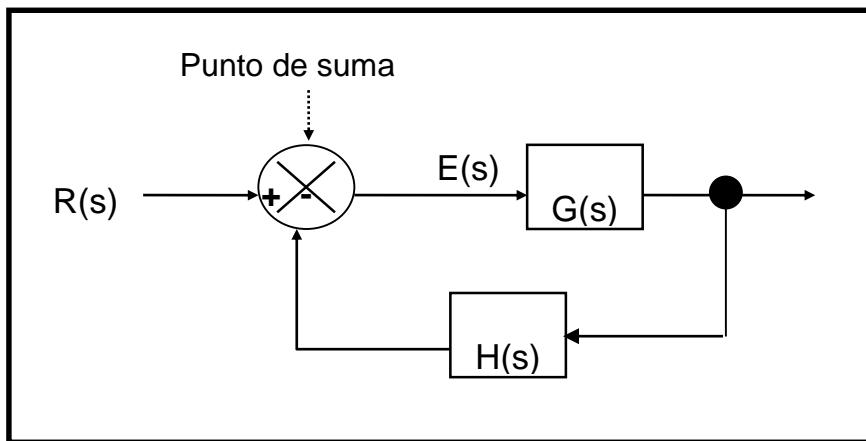


Figura 9.9: Sistema de Ciclo Cerrado

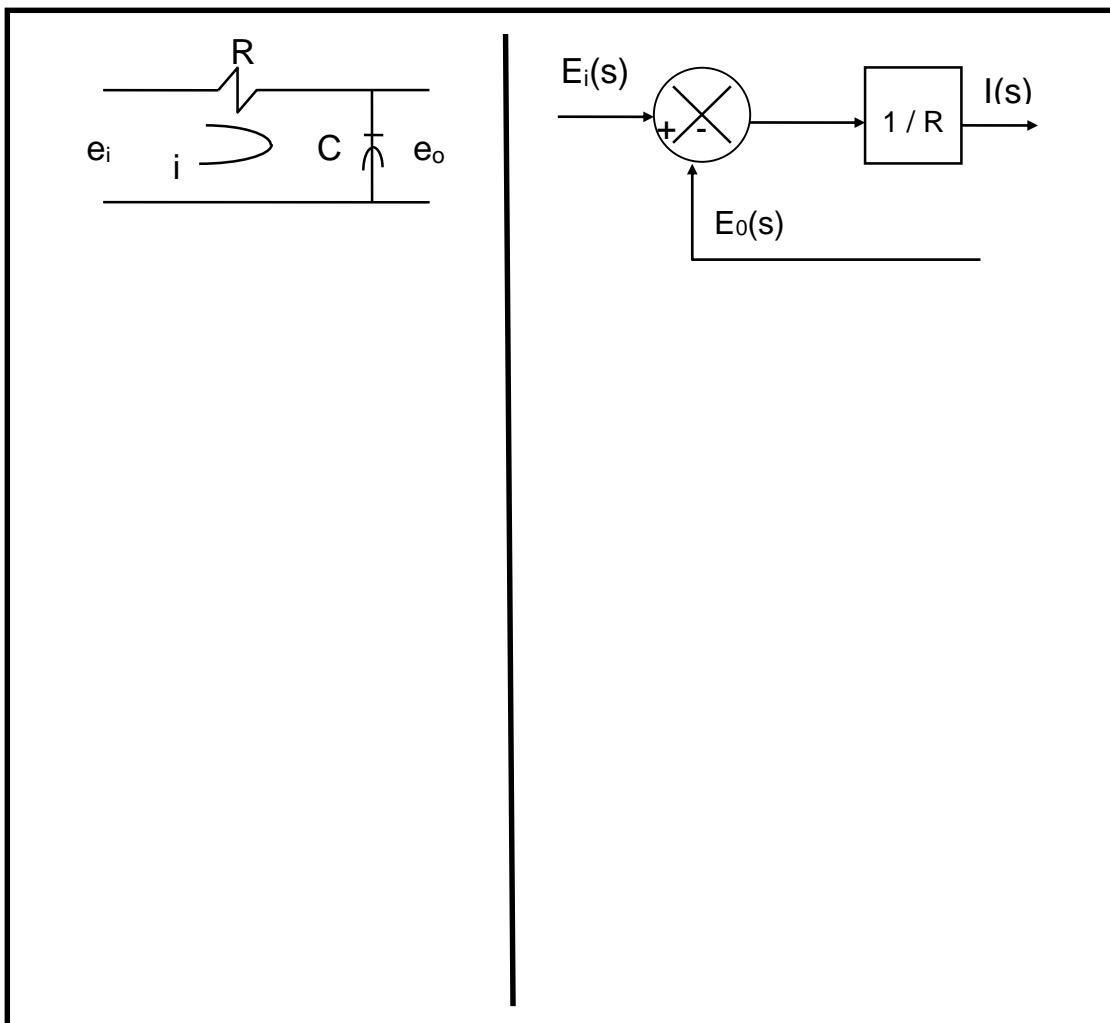
9.5.3 Realimentada o Directa

La relación de salida representada por: $\Theta_0(s)$ con respecto a la señal de error actuante $E(s)$ se llama función de transferencia pre alimentada, de modo que la función de transferencia pre alimentada será $= \Theta_0(s) / E(s) = G(s)$

Si la función de transferencia de la retroalimentación es unitaria, la función de transferencia pre alimentada y estas serán las mismas.

9.6 Procedimientos para trazar un Diagrama de Bloques

Para trazar un diagrama de bloque de un sistema, primero se escriben las ecuaciones que describen el comportamiento dinámico de cada componente, luego se toman las transformadas de Laplace de esas ecuaciones, suponiendo condiciones iniciales cero, y se representa, individualmente cada ecuación con la transformada de Laplace en forma de bloques completos. Como ejemplo, sea el circuito RC que se ve en la siguiente figura 9.10:



(a)

(b)

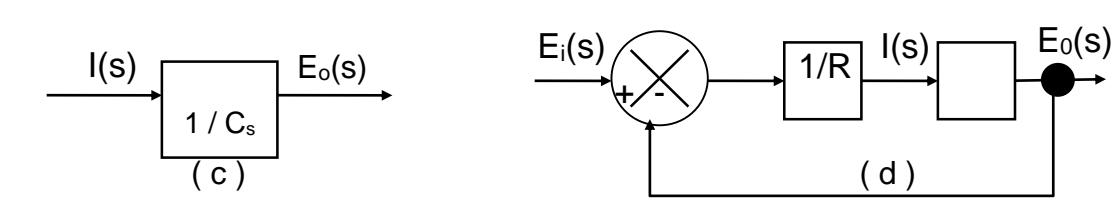


Figura (a) Circuito RC; (b) Diagrama de bloques que representa la ecuación (3); (c) Diagrama de bloques que representa la ecuación (4); (d) Diagrama de bloques del circuito RC

Figura 9.10: Diagramas de Bloques

Las ecuaciones de este circuito son:

$$i = e_i - e_o / R \quad (1)$$

$$e_o = dt / C \quad (2)$$

Las transformadas de Laplace de las ecuaciones 1 y 2, con condición inicial cero, son

$$I(s) = [E_i(s) - E_o(s)] / R \quad (3)$$

$$E_o(s) = I(s) / C_s \quad (4)$$

La ecuación 3 representa una operación de suma, y el diagrama correspondiente es el de la figura anterior. La ecuación 4 representa el bloque como puede verse en la figura anterior (c). Reuniendo estos dos elementos se obtiene el diagrama de bloque global del sistema que se ve en la figura anterior (d).

9.7 Reducción de un Diagrama de Bloque

Es importante notar que se pueden conectar los bloques en serie solamente si la salida de un bloque no es afectada por el bloque inmediato siguiente. Si hay cualquier efecto de carga entre los componentes, es necesario combinar esos componentes en un único bloque. Se puede representar con un único bloque cualquier cantidad de bloques en escala que representa componentes que no cargan, cuya función de transferencia es simplemente el producto de las funciones de transferencia individuales. Es posible simplificar un diagrama de bloque muy complejo con muchos ciclos de retroalimentación por una modificación paso a paso, utilizando reglas del álgebra de diagramas de bloques. Simplificando el diagrama de bloques por arreglos y sustituciones, se reduce notablemente la tarea a efectuar en el análisis matemático subsiguiente. Hay que notar, sin embargo, que, al simplificar el diagrama de bloques, los nuevos bloques se vuelven más complejos, debido a que se generan nuevos polos y ceros.

Al simplificar un diagrama de bloques, debe recordarse lo siguiente:

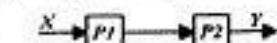
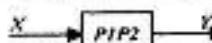
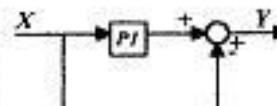
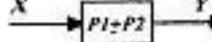
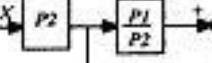
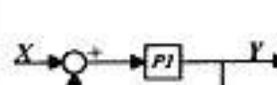
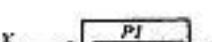
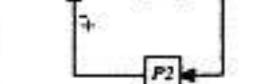
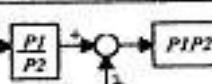
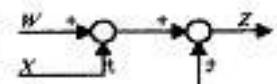
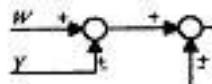
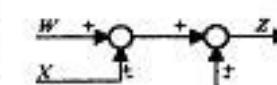
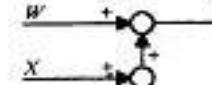
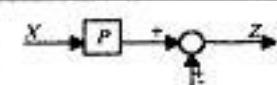
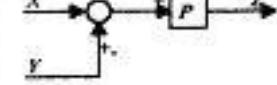
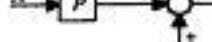
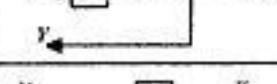
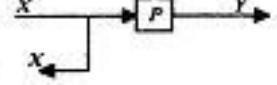
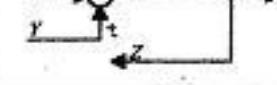
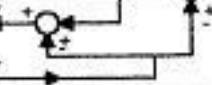
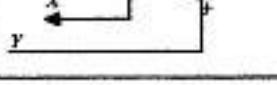
- El producto de las funciones de transferencia en la dirección de alimentación directa debe mantenerse constante.
- El producto de las funciones de transferencia alrededor del ciclo debe mantenerse constante.

Una regla general para simplificar un diagrama de bloque es desplazar los puntos de bifurcación y puntos de suma, intercambiar los puntos de suma y finalmente reducir los ciclos internos de retroalimentación.

9.8 Reglas del Álgebra de los Diagramas de Bloque

Los diagramas en bloque de sistemas de control complicados se pueden simplificar usando transformaciones que se pueden derivar fácilmente. La primera transformación importante, donde se combinan bloques en cascada se ha presentado. Con el fin de dar una visión completa de esta transformación que se incluye en la siguiente ilustración de los teoremas

de transformación. La letra P se utiliza para representar cualquier función de transferencia y W, X, Y, Z denotan cualquier señal en el dominio s.

TRANSFORMACIÓN		ECUACIÓN	DIAGRAMA EN BLOQUE	DIAGRAMA EN BLOQUE EQUIVALENTE
1	Combinación de bloques en Cascada	$Y = (P1P2)X$		
2	Combinación de bloques paralelos o eliminación de un lazo directo	$Y = P1X \pm P2X$		
3	Eliminación de un bloque de trayectoria directa	$Y = P1X \pm P2X$		
4	eliminación de un lazo de retroalimentación	$Y = P1(X \pm P2Y)$		
5	eliminación de un bloque de un lazo de retroalimentación	$Y = P1(X \pm P2Y)$		
6a	Redistribución de los puntos de suma	$Z = W \pm X \pm Y$		
6b	Redistribución de los puntos de suma	$Z = W \pm X \pm Y$		
7	Desplazamiento de un punto de suma hacia delante de un bloque	$Z = P(X \pm Y)$		
8	Desplazamiento de un punto de suma más allá de un bloque	$Z = P(X \pm Y)$		
9	Desplazamiento de un punto de toma hacia delante de un bloque	$Y = PX$		
10	Desplazamiento de un punto de toma más allá de un bloque	$Y = PX$		
11	Desplazamiento de un punto de toma hacia delante de uno de suma	$Z = X \pm Y$		
12	Desplazamiento de un punto de toma más allá de uno de suma	$Z = X \pm Y$		

A continuación, se presentan reglas comunes utilizadas en aritmética para los diagramas de bloques o de la transformada de Laplace, en forma escalar y matricial.

9.8.1 Reglas para Cantidades Escalares

$$A(s)B(s) = B(s)A(s)$$

$$A(s) + B(s) = B(s) + A(s)$$

$$A(s)[X(s) + Y(s)] = A(s)X(s) + A(s)Y(s)$$

$$A(s)[X(s) + Y(s)] = X(s)A(s) + Y(s)A(s)$$

$$[X(s) + Y(s)]A(s) = X(s)A(s) + Y(s)A(s)$$

$$C(s) = R(s) \frac{G(s)}{1 + G(s)H(s)} = \frac{G(s)}{1 + G(s)H(s)} R(s)^a$$

9.8.2 Reglas para Cantidades Matriciales

$$A(s)B(s) = B(s)A(s)$$

$$A(s) + B(s) = B(s) + A(s)$$

$$A(s)[X(s) + Y(s)] = A(s)X(s) + A(s)Y(s)$$

$$A(s)[X(s) + Y(s)] = X(s)A(s) + Y(s)A(s)$$

$$[X(s) + Y(s)]A(s) = X(s)A(s) + Y(s)A(s)$$

$$C(s) = [I + G(s)H(s)]^{-1} G(s)R(s) \text{ con términos en este orden sólo en este orden. } I = \text{matriz identidad.}$$

9.9 Usos y Beneficios

Los diagramas de bloques se utilizan para representar los diferentes componentes de un sistema de control, los cuales están conectados de tal manera que se caracterizan por su relación funcional dentro del sistema. Las flechas que interconectan los bloques en el diagrama son usadas para representar la dirección del flujo de la energía de control o información. Un diagrama de bloque permite visualizar más fácilmente el comportamiento dinámico de un sistema de control, describiendo las funciones de sus componentes y las relaciones que existen entre estos elementos. Es importante que los términos que se utilizan en un diagrama de bloque se entiendan claramente y se recuerden. Para representar las variables de entrada y salida de cada elemento se usan letras minúsculas, estas cantidades representan funciones de tiempo. Las letras mayúsculas representan funciones de Laplace de cantidades que son funciones de la variable compleja o transformada de Fourier de cantidades (funciones de frecuencia) que son funciones de la variable imaginaria pura. Las funciones generalmente se abrevian con una letra mayúscula que aparece sola. Las funciones de frecuencia nunca se abrevian.

Los diagramas de bloque poseen un amplio uso en diferentes aplicaciones de la vida diaria y del saber humano como los son:

- La Medicina
- Economía
- Biología
- Electricidad
- Hidráulica

Los cuales contribuyen y benefician en diferentes formas como:

- Salvar vidas, como es el caso de la Medicina.
- Diseñar un buen plan económico.
- Construir una planta eléctrica o hidráulica más eficiente.

9.10 Ejemplos

En el siguiente punto se presentará información que será de ayuda para describir los conceptos fundamentales de la teoría de control y se especificará el funcionamiento de dos ejemplos típicos:

- el *Acondicionador de Aire*
- el *Semáforo Semi-Inteligente*

que funcionan como *sistemas de ciclo cerrado de retroalimentación negativa*.

En el Acondicionador de Aire podemos describir brevemente cuáles son sus principales componentes, estos son: el elemento responsivo o regulador, este es responsable de hacer los cambios en la temperatura, presión, humedad o alguna otra característica que tenga que ser controlada. El control actuador, el cual inicia y para (o ajusta) la operación del equipo; son dispositivos tales como compuertas reguladoras, relevadores o paquetes electrónicos con sus circuitos correspondientes.

El equipo bajo control consta de equipo de calefacción o enfriamiento o ventiladores o bombas. La retroalimentación del equipo en sí mismo produce un resultado final haciendo operar al dispositivo regulador, por ejemplo, cuando en un cuarto se ha logrado la temperatura deseada las señales del regulador, cortan o reducen la capacidad del equipo de enfriamiento, porque han sido satisfechas las necesidades requeridas.

El semáforo *Semi-Inteligente* es un sistema dirigido por impulsos mediante sensores que dirigen el flujo vehicular, controlado por un regulador de tiempo que indica si hay más vehículos esperando pasar, este envía la señal a un panel o caja de control en donde se ejecutan los cambios de las luces dependiendo del orden de los vehículos.

9.10.1 Acondicionador de Aire

El acondicionador de aire es un sistema compuesto por diferentes elementos mecánicos, electrónicos y eléctricos que interactúan entre sí para realizar un trabajo, lo cual permite reducir la temperatura a través de la remoción de calor de espacios para mantener la confortabilidad de los mismos.

Acondicionar el aire es regular las condiciones ambientales, ya sea calefaccionando o refrigerando, de manera tal que la temperatura, humedad y grado de pureza del aire sean las adecuadas. Por ejemplo, se acondiciona el aire para realizar algunos procesos industriales, en las salas de terapia intensiva de los hospitales, en los centros de computos, etc.

El objetivo del acondicionador de aire es el de encargarse de producir frío o calor y de impulsar el aire tratado a la vivienda o local.

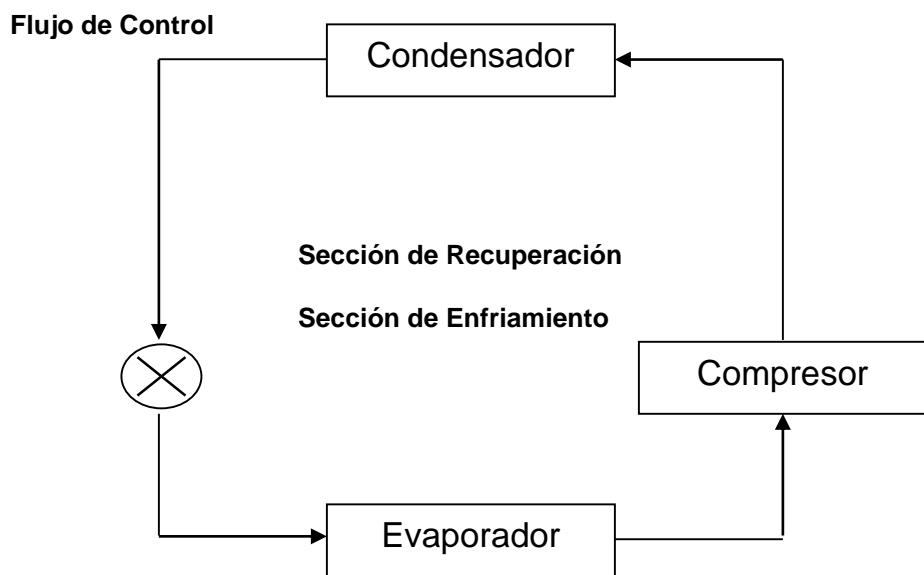


Figura 9.11: Diagrama de Bloque del Acondicionador de Aire

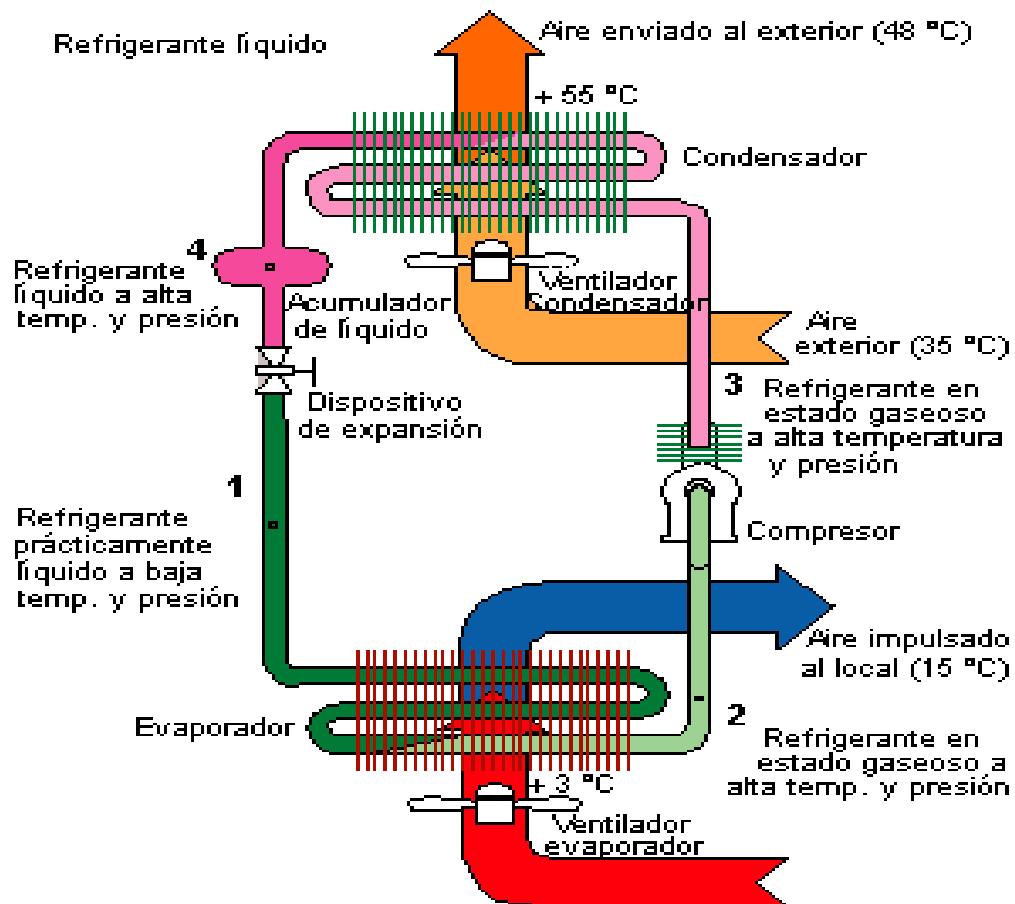


Figura 9.12: Esquema del Circuito Frigorífico (Generalmente, los acondicionadores de aire funcionan según un ciclo frigorífico similar al de los frigoríficos y congeladores domésticos).

¿Cómo funciona un Acondicionador de Aire?

El acondicionador de aire toma aire del interior de una recamara pasando por tubos que están a baja temperatura estos están enfriados por medio de un líquido que a su vez se enfriá por medio del condensador, parte del aire se devuelve a una temperatura menor y parte sale impulsada por el panel trasero del aparato, el termómetro está en el panel frontal para que cuando pase el aire calcule la temperatura a la que está el ambiente dentro de la recamara, y así regulando que tan frío y que tanto debe trabajar el compresor y el condensador.

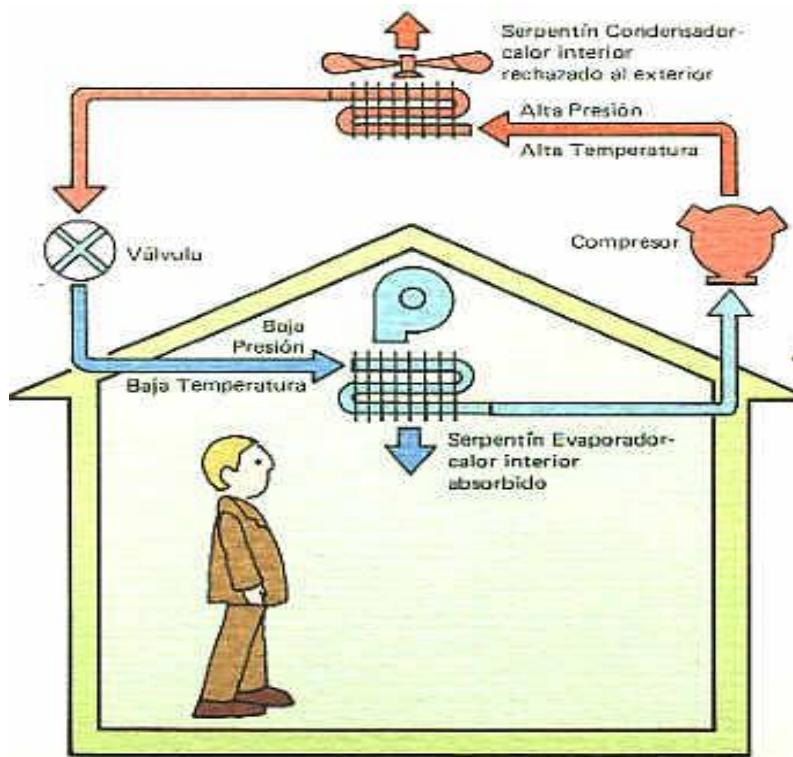


Figura 9.13: Diagrama de un Acondicionador de Aire Típico

En la figura se muestra un acondicionador de ventana tradicional, funcionando en condiciones típicas.

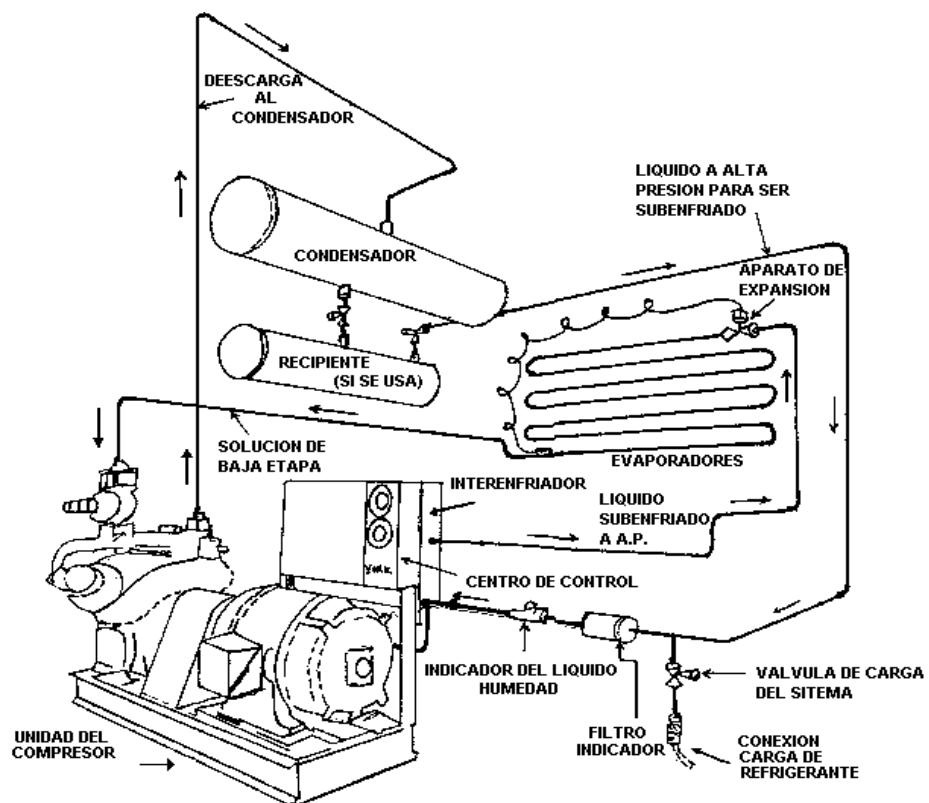


Figura 9.14: Dibujo Técnico del Sistema Refrigerante del Acondicionador de Aire

Descripción Funcional del Enfriamiento de la Refrigeradora Doméstica:

Las refrigeradoras domésticas utilizan para enfriar, un sistema de enfriamiento por compresión de vapor.

El ciclo de refrigeración por compresión de vapor tiene cuatro dispositivos básicos que son:

- Dispositivo de expansión (dispositivo de control de flujo)
- Evaporador
- Compresor
- Condensador

El proceso A-B a través del dispositivo de control de flujo:

El refrigerante líquido entra al dispositivo de expansión en el punto A, a una presión y a una temperatura relativamente alta. Existen diversos dispositivos de expansión, en el caso de las refrigeradoras domésticas, se utiliza un dispositivo llamado tubo capilar. Este dispositivo de control de flujo tiene una abertura estrecha, lo que da como resultado una

gran pérdida de presión al fluir el refrigerante a través del mismo. El refrigerante sale por el punto B a una presión muy inferior comparada con la presión con la que entró. Debido a que esta presión es inferior a la presión de saturación correspondiente a la temperatura con que entró, una parte del líquido refrigerante se evaporiza inmediatamente y rápidamente a gas. La porción del líquido que se evapora toma el calor latente necesario para su evaporación de la mezcla que fluye, enfriándola de esta manera. El refrigerante sale de la válvula como una mezcla de líquido y vapor en estado saturado.

El proceso B-C a través del Evaporador:

El refrigerante fluye a través de la tubería del evaporador, de B a C. La sustancia que se debe enfriar (aire), fluye por el exterior de los tubos. Se halla a una temperatura más elevada que la del refrigerante dentro del evaporador. Por consiguiente, el calor fluye de la sustancia al refrigerante, a través de la pared del tubo y el aire se enfriá. Debido a que el refrigerante líquido dentro del evaporador ya se encuentra en su temperatura de saturación (su punto de ebullición), el calor que gana hace que se evapore al pasar por el evaporador. Por lo general, el refrigerante sale del evaporador ya sea como un vapor saturado o un vapor sobrecalentado.

El proceso C-D a través del compresor:

El compresor hace entrar al vapor por el lado de la succión, para luego comprimirlo a una presión elevada, adecuada para efectuar la condensación. Esta presión es ligeramente mayor a la presión con la cual entró al dispositivo de control de flujo. Se requiere trabajo para comprimir el gas; este trabajo procede de un motor o una máquina que mueve el compresor. Este trabajo contribuye a aumentar la energía almacenada del vapor comprimido, resultando en un aumento en la temperatura.

El proceso D a través del condensador:

El gas a alta presión que descarga el compresor fluye a través de la tubería del condensador, de D a A. Un fluido, tal como el aire o el agua, fluye por el exterior de la tubería. El calor fluye a través de las paredes del tubo, desde el refrigerante a mayor temperatura hasta el fluido de enfriamiento. Como el refrigerante está sobrecalentado cuando entra al condensador, primero se enfriá hasta que alcanza su temperatura de saturación. La remoción adicional de calor resulta en la condensación gradual del refrigerante, hasta que se licua en su totalidad. El refrigerante puede salir del condensador como un líquido saturado o sub-enfriado.

¿Qué Parámetros controla el Equipo de Acondicionador de Aire?

Temperatura

Es evidente, pero también muy importante: el acondicionador de aire elimina el calor del aire, por lo que el termómetro señala una temperatura más fresca.

Humedad

El aire de nuestro entorno natural contiene humedad en forma de vapor, y el nivel saludable de humedad relativa del aire se sitúa entre el 40-60%. La capacidad del aire de contener vapor de agua depende sobremanera de su temperatura: el aire cálido puede atrapar más humedad que el fresco. Como el aire cálido atraviesa las baterías frías del sistema de acondicionador de aire, el exceso de humedad desaparece del aire al condensarse en las baterías, recogiéndose en forma de líquido en el sistema.

Pureza del Aire

El aire está repleto de partículas diminutas: polvo, polen, humo, contaminación, bacterias, esporas y otras muchas impurezas. El sistema de acondicionador de aire está dotado de un filtro que retiene tales partículas y devuelve al interior un aire depurado y más saludable.

Circulación del Aire

Todos estos efectos beneficiosos no se repartirían de manera uniforme por los recintos si el sistema no renovase el aire de forma óptima. El ventilador y las rejillas del sistema deben ajustarse de forma que hagan circular el aire de manera uniforme en el lugar de instalación, pero sin expulsarlo directamente hacia las personas situadas junto a ellos.

Tabla de Descripción de Entidades:

Compresor:	Es el primer elemento que se analiza, en los cuatro componentes de un sistema de compresión de vapor (compresor-condensadores-dispositivos de expansión-evaporador). Es el corazón del sistema de compresión de vapor. Es el elemento que se encarga de hacer circular el gas refrigerante a través de las tuberías del ciclo de la refrigeración.
Condensador:	Los condensadores reciben el vapor refrigerante recalentado procedente del compresor, eliminan el recalentamiento del vapor y, a continuación, lo licuan (convertir en líquido). El condensador es el

	lugar donde se produce la eliminación de calor de un sistema de refrigeración.
Condensación:	Proceso en el cual el gas o vapor es transformado a líquido.
Dispositivo de Expansión (Válvula Capilar):	Después del condensador y el compresor, el elemento fundamental que sigue en el sistema de compresión de vapor es el dispositivo de expansión. El propósito de dispositivo de expansión es doble, debe reducir la presión del líquido refrigerante y debe regular el paso del refrigerante al evaporador. Es un sistema de control de refrigerante. Usualmente consiste en algunos centímetros de un tubo con un pequeño diámetro interno. La fricción del refrigerante líquido y las burbujas del refrigerante evaporizado dentro del tubo, sirven para restringir el flujo de tal manera que las presiones de alta y de baja presión se mantienen mientras el compresor funciona. Un sistema de tubos capilar permite que los lados de alta y baja se nivelen cuando el equipo está apagado. Además, un tubo de diámetro pequeño es usado para conectar bulbos de control de temperatura con mecanismo de control.
Estado de Saturación:	Estado de ebullición, temperatura en la cual la presión de vapor de un líquido es igual a la presión externa.
Evaporador:	Un evaporador de un sistema de refrigeración es un intercambiador de calor en el cual el calor pasa desde una sustancia que se va a enfriar hasta el refrigerante en ebullición. El propósito de un sistema de refrigeración es absorber el calor del aire, agua, o cualquier otra sustancia. Esta función la realiza el evaporador.
Refrigerante:	Sustancia usada en los mecanismos de refrigeración para absorber calor en el serpentín del evaporador cambiando de estado de líquido a gas y liberándolo en un condensador mientras la sustancia cambia de nuevo de gas a líquido.
Refrigerante Vaporizado:	Refrigerante en estado de vapor.
Temperatura Ambiente:	Es evidente, pero también muy importante: el acondicionador de aire elimina el calor del aire, por lo que el termómetro señala una temperatura más fresca.
Transferencia de Calor:	Movimiento de calor de un cuerpo o sustancia a otro. El calor puede ser transferido por radiación, conducción, convección o una combinación de estos tres métodos.
Tubería:	Línea de transporte de fluido que tiene una pared delgada.
Vapor:	La fase gaseosa de una sustancia que, bajo condiciones ordinarias, existe como un líquido o un sólido.

9.10.2 Semáforo Semi – Inteligente

Consideremos al semáforo semi-inteligente un sistema dinámico ya que en él se aplican principios y técnicas de la teoría de control de retroalimentación.

El semáforo semi-inteligente está diseñado bajo los principios de la implementación de cola. Una cola es un grupo ordenado de elementos los cuales se añaden por un extremo y se suprime por el otro extremo.

El funcionamiento del semáforo semi-inteligente depende de la total sincronización del conjunto de sus componentes, hemos considerado que dichos componentes son:

- el semáforo,
- los cables sensores y
- la caja de control de ESPERA-PASO.

La comunicación entre estos tres componentes se hace por medio de señales (bit o impulsos) que viajan a través de cables, de acuerdo a la señal recibida o transmitida se ejecutará una acción en particular.

A continuación, describiremos cómo funciona el sistema, tomando como referencia el estado de espera:

Cuando un vehículo acciona los cables sensores, una señal es enviada a la cajilla de control ESPERA-PASO de forma tal que se vayan registrando y programando para determinar la prioridad del carril de donde se ha hecho la solicitud de paso vehicular. Por ejemplo, tomando como referencia los puntos cardinales: el carril norte, luego el carril sur, luego el carril oeste y nuevamente al carril norte es enviado una señal; esta es considerada en el orden en que se halla registrado cada llegada. Cuando los vehículos tengan el acceso al paso, inmediatamente, se activan los relojes, que en nuestros ejemplos lo llamamos reloj principal y reloj auxiliar. El reloj principal tiene como función registrar un tiempo máximo en el cual los vehículos puedan circular. Al finalizar el tiempo máximo (posiblemente 2 minutos) el reloj principal emite una señal a los controles para que se indique que el período de flujo vehicular ha terminado (luz verde) y se da la orden para que los vehículos se detengan (luz roja). Si al cabo de transcurrir el tiempo máximo, existen más vehículos en el mismo carril, ellos tomarán el estado de espera y se acogerán al ciclo de registro de solicitud de paso siempre y cuando se compruebe la existencia de una solicitud de paso en otro carril. La función del reloj auxiliar es la de comprobar la continuidad del fluido. Cada vez que un vehículo haga contacto con los cables sensores este reloj es reiniciado nuevamente a cero. Si se cumple su tiempo

máximo, este emitirá una señal a los controles para indicar que se ha roto la continuidad del flujo y por lo tanto no es necesario que el tiempo máximo del reloj principal se complete para dar la autorización del cambio de luz de VERDE a ROJA, permitiendo entonces el avance vehicular de la siguiente solicitud de paso (véase la figura 9.13).

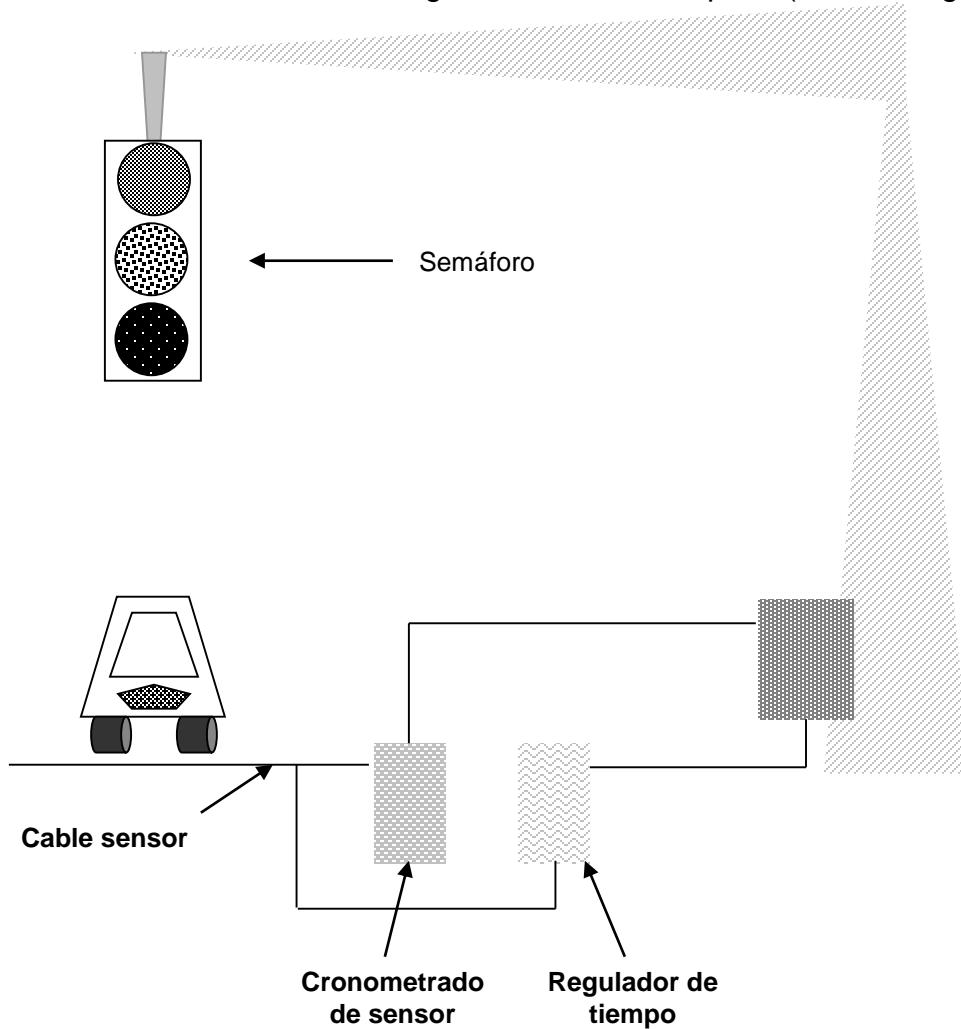


Figura 9.15: Diagrama Técnico del Semáforo Semi-Inteligente

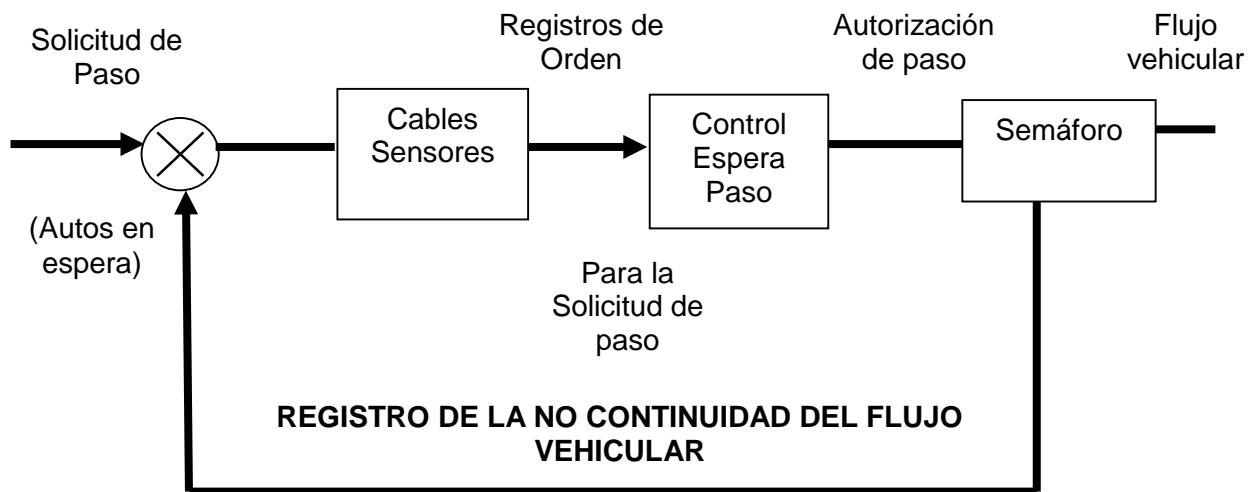


Figura 9.16: Diagrama de Bloque del Sistema de Control de un Semáforo Semi-Inteligente

Los componentes básicos necesarios para establecer un sistema de semáforos semi-inteligentes son los siguientes:

Semáforo:	Unidad en la cual se desplazan los colores que determinan la espera o paso vehicular. Por reglas universales los colores son el verde, amarillo y rojo, y la función de cada uno de ellos es la siguiente: <ul style="list-style-type: none"> • <u>Verde</u>: Indicador que autoriza al conductor del vehículo en espera el avance. • <u>Amarillo</u>: Indicador que le señala al conductor del vehículo que se hará un cambio de luz y debe tomar precauciones. • <u>Roja</u>: Indicador que autoriza al conductor del vehículo que debe detenerse para darle paso a otro carril en espera.
Cables Sensores:	Unidad diseñada y construida para percibir el contacto que indica que en el carril existen vehículos en espera o que se ha roto la continuidad del flujo. En ambos casos los cables sensores emiten una señal a la caja de control de ESPERA-PASO para que determine la acción a tomar. Por lo general estos cables están distribuidos por debajo del pavimento que recurre a cada carril a una distancia razonable entre uno y otro de forma tal que se asegure la existencia o no de vehículos en espera.

Caja de control Espera-Paso:	Bajo la implementación de programación esta unidad está diseñada para recibir todas las señales emitidas por los cables Sensores, procesarlos y luego generar señales que autoricen el avance de los vehículos en espera o la orden de espera a los vehículos, mediante el uso del semáforo (véase la figura 9.14).
------------------------------	---

9.11 Resumen

En este capítulo se han presentado los conceptos básicos de los diagramas de bloques los cuales se utilizan para entender el comportamiento conceptual y matemático de sistemas primordialmente electromecánicos. Los diagramas de bloques complejos se pueden simplificar con reglas sencillas, las cuales se derivan fácilmente. La letra F se utiliza para representar cualquier función de transferencia y W, X, Y, Z denotan cualquier señal en el dominio s. Los conceptos fundamentales de la teoría de control se presentan en dos ejemplos típicos: El Acondicionador de Aire y el Semáforo Semi-Inteligente.

Capítulo 11

Workflow

Capítulo 11

Workflow

11.1 Introducción

Hacer más con menos es más que un buen consejo para los negocios de hoy, es una receta para mejorar la calidad y mantener competitividad. Para hacerlo, los negocios deben re-pensar los procesos en todos los frentes - estratégica, organizativa y tecnológicamente - en una forma sinérgica.

Ya sea que se le llame transformación de los negocios o re-estructuración de los procesos de negocio, esto significa que las organizaciones deben re-diseñar los procesos críticos y capitalizarse en tecnología.

En el momento en que las organizaciones re-diseñen los procesos, deben continuar para auto modificarse y reflejar los cambios que se producen en el ambiente en que se desenvuelven. Responder rápida y eficientemente es una ventaja competitiva.

El Workflow se constituye entonces en una herramienta que facilita el seguimiento de los procesos, y automatiza el proceso rutinario y necesario.

11.2 ¿Qué es el Workflow?

El Workflow es una herramienta para automatizar y agilizar los procesos. La esencia del proceso Workflow es que un documento tiene que movilizarse de un lugar a otro, realizar un trámite con dicho formulario electrónico y continuar su proceso hasta su culminación.

A diferencia de los sistemas tradicionales cuya estructura es funcional (presentado por módulos. Ejemplo: Contabilidad, Planillas, Facturación, etc.), cuya operación estaba basada en el "registro" de documentos y/o operaciones para luego efectuar su posterior "procesamiento", el Workflow estará basado en uno de los más modernos conceptos de manejo empresarial: **"Administración sin papeles"** por lo que en el diseño del software se contempla la definición de procesos o flujos de trabajo denominados **"Workflow"** que comprenden una lista de actividades o eventos relacionados en forma lógica y secuencial; para la ejecución de cada una de éstas se podrá definir "aprobaciones electrónicas" (firmas electrónicas), reduciendo al mínimo la emisión física de documentos para visto bueno y firma, antes de continuar con el siguiente paso en la cadena administrativa.

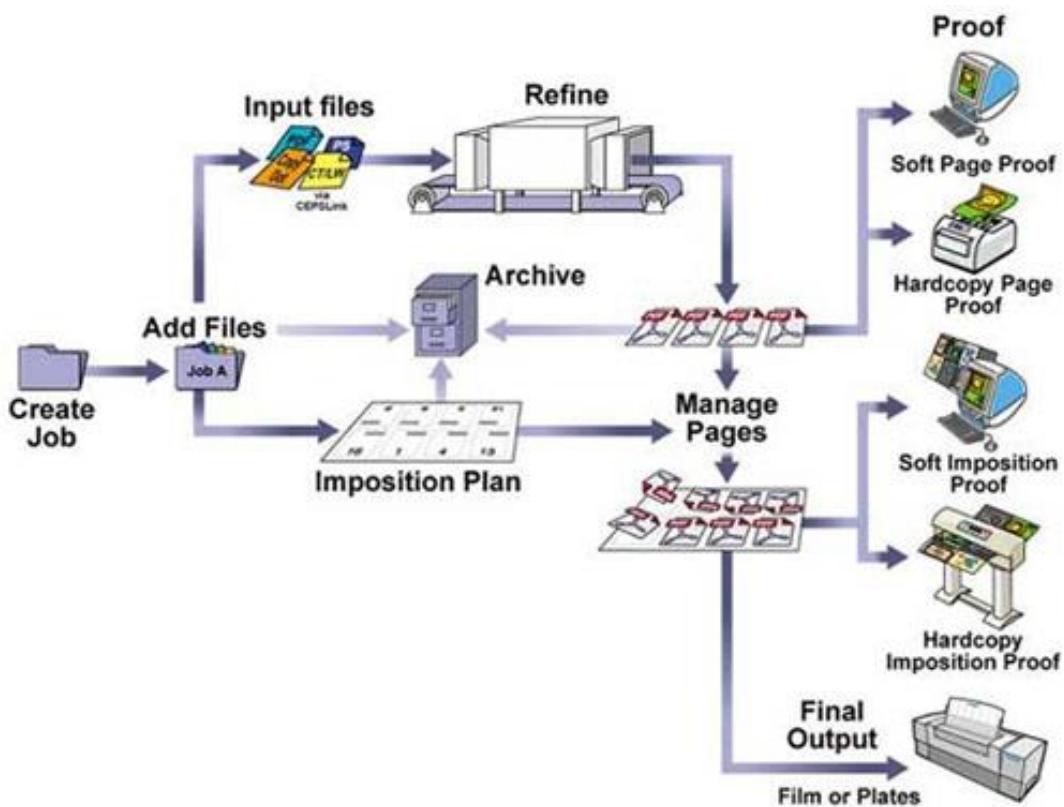


Figura 11.1: Ejemplo 1 de Workflow

Dicha situación generaba tiempos "muertos" o de espera en el traslado físico del documento desde el escritorio de la persona que generó el documento hacia el escritorio de la persona que autoriza y finalmente hacia la persona que debe de ejecutar o comunicar a la persona o entidad que requiere el documento. Por tal razón, el **tiempo total de la transacción** es siempre mucho mayor que la sumatoria de **los tiempos individuales** empleados para completar cada una de las etapas de trabajo.

A grandes rasgos, el flujo de trabajo podría definirse como un conjunto de mecanismos que automatizan el flujo de tareas dentro de una organización. Estos mecanismos relacionan entre sí los diferentes aspectos de la gestión, priorizan las actividades de cada empleado y optimizan las comunicaciones entre las distintas unidades operativas. Una aplicación de flujo de trabajo automatiza la secuencia de acciones, actividades o tareas utilizadas para la ejecución del proceso, incluyendo el seguimiento del estado de cada paso del mismo y la aportación de las herramientas necesarias para gestionarlo. Las herramientas de flujo de trabajo soportan encaminamiento estructurado y revisión y seguimiento de documentos, formularios y otras informaciones.

En conclusión, definiremos flujo de trabajo como: "cualquier tarea ejecutada en series o en paralelo por dos o más miembros de un grupo de trabajo para llegar a una meta común."

A continuación detallaremos esta definición:

Cualquier tarea: implica que workflow se refiere a un muy amplio campo de acción relacionado con actividades propias de empresas.

En serie o en paralelo: implica que los pasos en la tarea pueden ser ejecutados uno tras el otro o simultáneamente por diferentes personas, o una combinación de ambos.

Dos o más miembros: implica que si solo una persona ejecuta esa tarea no se considera workflow. Como la palabra workflow sugiere, una tarea es un flujo de trabajo cuando fluye de una persona a otra.

Meta Común: Individuos participando en workflow deben trabajar hacia una meta común. Si ellos trabajan en proyectos independientes no es considerado como workflow.

Dada la definición de workflow hay un gran número de actividades empresariales en una organización que se puede incluir en la categoría de un flujo de trabajo. Estas incluyen:

- Ordenes de compra
- Solicitud de Capital
- Revisión del rendimiento del personal
- Reportes semanales
- Aprobación de préstamos
- Procesamiento de quejas

11.3 ¿Cómo identifico la Necesidad de utilizar Workflow?

El siguiente cuestionario puede ayudarnos a identificar posibles lugares con alto potencial para desarrollar proyectos de imágenes y workflow.

- ¿Hay una cantidad significativa de papeles?
- ¿Hay personal dedicado exclusivamente a manejar papeles?
- ¿Es el papel usado para tomar decisiones o llevar adelante procesos?
- ¿Hay una pesada carga de trabajo generada por archivo -recuperación - archivo?
- ¿Más de un usuario necesita el mismo papel?
- ¿Lo necesitan al mismo tiempo?
- ¿Se hacen múltiples fotocopias regularmente?
- ¿Se necesitan los papeles en más de un lugar?
- ¿Se pierden frecuentemente los papeles?
- ¿El trabajo es impactado significativamente por el manejo de papeles?

Si la respuesta a varias de estas preguntas es **SI**, entonces estaremos en presencia de un potencial proyecto de flujo de proceso

El flujo de trabajo o workflow supone hoy una tecnología en alza, directamente relacionada con la gestión de imágenes y documentos. Su importancia dentro de los procesos de reingeniería de negocios que hoy afrontan las empresas, le está convirtiendo además en una premisa básica de la agilización y descentralización de las actividades administrativas y comerciales impuestas por las nuevas tendencias que rigen en las organizaciones.

A la vez que crece la presencia de los sistemas de flujo de trabajo en las empresas y aumentan su funcionalidad y prestaciones, se van difuminando las líneas divisorias que separan este tipo de aplicaciones de las imágenes (imaging) y gestión de documentos, proyectos y bases de datos. De hecho, muchos productos de flujo de trabajo acceden a ficheros basados en host o en servidor, sistemas de gestión de bases de datos y sistemas de gestión de documentos. A veces, incluso, todos los ficheros relacionados con una tarea particular pueden ser encaminados como una carpeta conjunta y presentados a los usuarios destinatarios como una lista de objetos pendientes. Asimismo, múltiples versiones de un documento particular pueden ser seguidas por sistemas de gestión de documentos basados en servidor y, una vez culminado el proceso, los contenidos de un formulario pueden ser escritos para bases de datos corporativas compartidas.

Esta integración conceptual y tecnológica puede con el tiempo, llegar tan lejos como para que el propio concepto de flujo de trabajo acabe desapareciendo como entidad individualizada, pasando a identificar un componente más de determinados tipos de aplicaciones. Pero lo cierto es que, hoy en día, los sistemas de flujo de trabajo constituyen un segmento bien diferenciado del mercado de trabajo en grupo o groupware, con soluciones comerciales y aspectos tecnológicos particulares.

11.4 Conceptos Básicos de Workflow

Existen cuatro conceptos básicos en materia de Workflow:

- Lógica de procesos: Concepto fundamental del cual se obtiene la representación de la definición de cada proceso con una metodología predeterminada. Se mantiene un seguimiento del estado de cada instancia a medida que se progresó en la tarea y se empuja el proceso hacia la siguiente etapa de acuerdo a la lógica que se le ha definido.
- Correcta concordancia entre personas y tareas. Los sistemas workflow ayudan a asegurar que la tarea que se necesita efectuar es hecha por la persona indicada, gracias a que generalmente están basados sobre sistemas de mensajería robustos.
- Entrega de recursos de información para las tareas. Cuando los recursos de información son basados en computadoras, los sistemas Workflow pueden asegurar que las tareas que necesitan ejecutarse tienen la información necesaria para ser completadas.

- Administración de procesos. Concepto clave, debido a que las organizaciones están bajo constante presión para mejorar el uso de sus recursos. Para los encargados de sistemas, la capacidad para administrar esos procesos es más crítica que la capacidad para construirlo en forma eficiente.

Los sistemas Workflow tienen fortalezas obvias en el control de los procesos gracias a su soporte automatizado, también prometen la ayuda a la administración al hacer que los procesos sean lógicos y explícitos en discretas capas de representación del diseño y permitiendo a los diseñadores crear, agrupar y evaluar métricas relativas al tiempo, costos o calidad en el desempeño de las tareas constituyentes de dichos procesos.

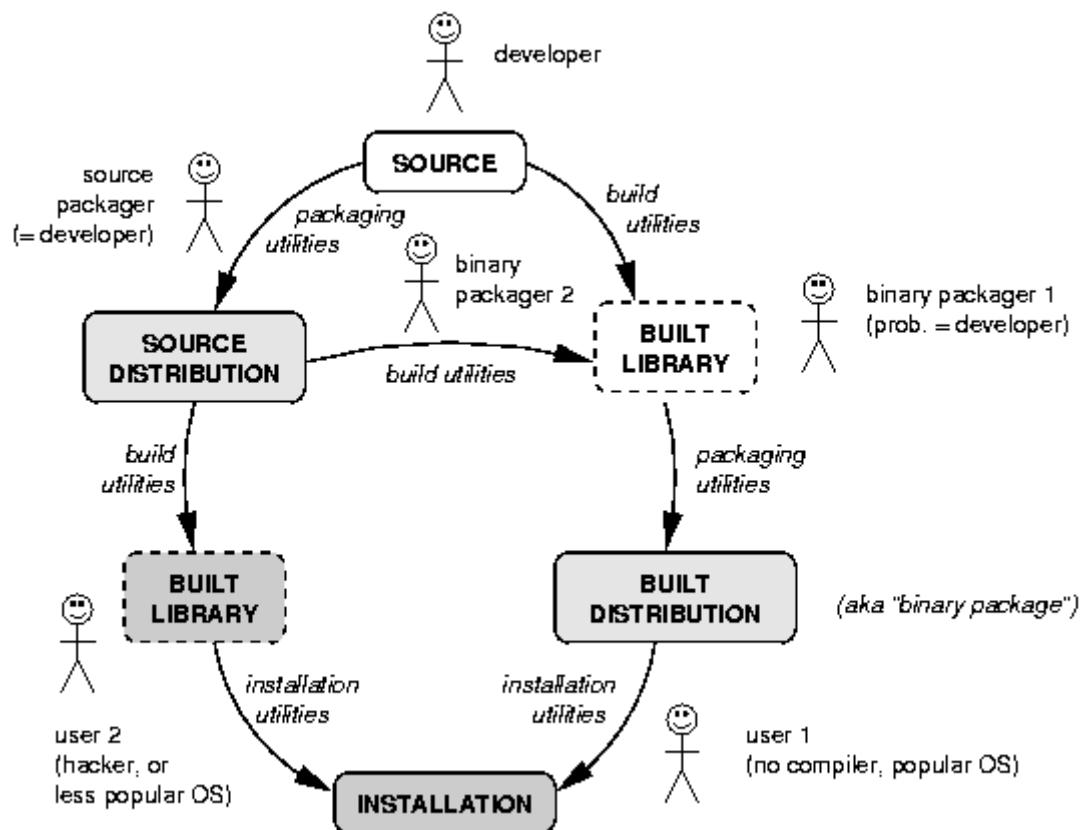


Figura 11.2: Ejemplo 2 de Workflow

En resumen podemos ver al workflow como un conjunto de métodos y tecnologías que nos ofrece las facilidades para modelar y gestionar los diversos procesos que ocurren dentro de una empresa. El workflow es el último, de una gran línea de facilidades propuestas en respuesta de las exigencias de las organizaciones. Las cuales apuntan a poder reaccionar tan rápido como sea posible ante la frenética demanda de la competencia.

Para ser clasificado como una solución automatizada workflow, una aplicación debe tener los siguientes ingredientes básicos.

- Un medio para diseñar mapas de workflow, preferiblemente gráficos.
- La habilidad de diseñar los formularios electrónicos para cada paso en el mapa del flujo de trabajo.
- La habilidad de incorporar aplicaciones externas como parte de la automatización del workflow.
- La habilidad de enlazar los formularios electrónicos a las bases de datos de la empresa.
- La habilidad de diseñar mapas de workflow con caminos o rutas condicionales y funciones basadas en reglas complejas de la empresa, preferiblemente sin programar.
- La habilidad de enviar la información recogida en cada paso hacia los pasos siguientes basados en funciones de trabajo, nombre de usuarios, o relaciones de reportes.
- La habilidad de monitorear el estatus del flujo de trabajo.
- La habilidad de medir el flujo de trabajo.
- La habilidad de estimular y probar el comportamiento del workflow.
- Una aplicación workflow debe respaldar muchos usuarios y ser extremadamente confiable y segura.
- El flujo de trabajo debe ser accesible internamente y sobre el Internet y a través de numerosas plataformas.

11.5 El Workflow o Flujo de Procesos y la Reingeniería

Uno de los aspectos clave de la importancia del flujo de trabajo es su íntima implicación en la pujante reingeniería de procesos de negocio, por cuanto representa un cambio profundo del modo de llevar a cabo las tareas comerciales y administrativas de las empresas. Entre otras características, las aplicaciones de flujo de trabajo se adaptan perfectamente a las organizaciones con estructuras horizontales, cada vez más presentes en todos los ámbitos empresariales, aportando la flexibilidad necesaria para eliminar o, al menos, reducir la burocracia y potenciar la capacidad de decisión de cada empleado. Asimismo, es un hecho que actualmente las organizaciones desean obtener un mayor y mejor control de la información, facilitando su acceso y su eliminación de acuerdo con esquemas definidos por ellas mismas.

Un factor que está contribuyendo, además, a la difusión de estos sistemas viene determinado en gran medida por los propios avances que se están produciendo en las infraestructuras de redes de las empresas, cuya disponibilidad y ancho de banda crece rápidamente, y en la gran generalización de los PCs como herramienta de trabajo. Ambos fenómenos hacen posible adoptar aplicaciones de flujo de trabajo sin necesidad de efectuar costosas inversiones adicionales.

En general, podría decirse que el flujo de trabajo está siendo implementado en cuatro tipos de aplicaciones:

- proceso de imágenes,
- proceso de formularios,
- reglas y filtrado de correo electrónico y
- aplicaciones de bases de datos compartidas.

11.6 Elementos del Flujo de Trabajo

Los seis principales elementos de un proceso de flujo de trabajo son:

11.6.1 Rutas

Definen el orden en el que han de fluir los objetos (documentos, formularios, datos, aplicaciones...), de modo que puedan ser capaces de moverse a través de diferentes rutas secuenciales y recomponerse después, en un punto dado, en una única ruta.

11.6.2 Los Objetos

Deben ser enviados tanto en modalidad de transmisión o en modalidad de difusión (broadcast), y en cualquier orden descrito por el usuario.

11.6.3 Reglas

Las reglas determinan qué información debe dirigirse a través de la ruta y a quién, función conocida como encaminamiento condicional o tratamiento de situaciones de excepción. En ocasiones, los productos de flujo de trabajo basan la definición de reglas en scripting u otras actividades de programación. Otros productos, especialmente los de correo electrónico, proporcionan motores de reglas que permiten a los usuarios sin conocimientos técnicos escribir reglas relativamente complejas mediante opciones de menús y listas de "arrastrar y soltar".

11.6.4 Roles

En un proceso de flujo de trabajo, los roles fijan las funciones a realizar con independencia de quién las lleve a cabo. Es necesario disponer de tablas de roles (generalmente una tabla o una base de datos en la que se establecen relaciones de asociación entre usuarios y roles determinados), en la que se puedan introducir cambios de una sola vez y fijar sustitutos para cada rol.

11.6.5 Procesos

Los procesos comprenden la realización de una serie de pasos en base a determinadas condiciones comerciales. Los procesos establecidos, que controlan y dirigen la actividad de una empresa, son tan variados y tan personales como los usuarios que participan en ellos. Con frecuencia, los procesos no son diseñados, sino que se identifican a posteriori y se extraen del uso o actividad común. Con frecuencia se utilizan herramientas y

metodologías de diseño de procesos comerciales para definir los procesos, que pueden abarcar varias aplicaciones.

11.6.6 Prácticas

Las prácticas son aquellas cosas que suceden en las empresas y pueden considerarse como actos que “infringen las reglas” y que hacen que el proceso funcione realmente. Con demasiada frecuencia, al diseñar un flujo de trabajo no se tiene en cuenta la práctica del mundo real que describe cómo ha evolucionado el proceso. Sólo capturando las prácticas relevantes es posible automatizar la forma en que se realiza la actividad comercial.

11.7 Pasos para Construir un Workflow

Un flujo de trabajo es creado con los siguientes pasos:

- Definir una actividad o tarea que un grupo de trabajo necesite desarrollar y las reglas que rigen esa actividad.
- Dividir la tarea en “sub-tareas” o “sub-procesos” a los cuales llamaremos **pasos**. Cada paso representa una lista bien definida de cosas que deben ser ejecutadas por un individuo y que deben ser ejecutados en orden lógico. Una tarea se puede dividir en pasos de diferentes maneras. Es aquí donde el juicio empresarial es requerido para decidir si hay que dividir una tarea en diferentes pasos y donde debe hacerlo.
- Decidir las habilidades requeridas para ejecutar cada paso. Debe especificar las funciones del trabajo y/o los individuos que se deben localizar para ejecutar los pasos.
- Identificar todas las aplicaciones externas usadas para completar la tarea y determinar si hay manuales de pasos que se pueden automatizar como una parte del workflow.
- Decidir la secuencia en que los pasos deben ser ejecutados.
- Si alguno de los pasos son ejecutados con una base condicional o sea que deben cumplir ciertas condiciones para ser ejecutadas, se debe identificar esos pasos y definir las condiciones.
- Trazar un “**mapa**” del workflow que identifique los pasos y la secuencia, o el *flujo* en que los pasos deben ser ejecutados. Asociar las funciones de trabajo o individuales en cada paso.
- Crear las formas, documentos e instrucciones que deben ser usados por las personas en cada paso para ejecutar o desempeñar cada sub-tarea o sub-proceso.

11.8 Algunos de los Beneficios del Workflow

- Los procedimientos son formalizados, documentados y seguidos exactamente; se asegura que el trabajo se realice de la forma como fue planeada por la administración;
- Mejora la eficiencia del negocio. La automatización de diferentes procesos del negocio conlleva a la eliminación de los pasos innecesarios, y también se elimina el uso de papeles o documentos innecesarios;
- Se mejora la administración de los procesos, a través de la estandarización de los métodos de trabajo, enfoque que lleva los procesos del negocio a la simplificación de las tareas y racionalización de los recursos;
- Mejoramiento del servicio al cliente, interno y externo, con procesos consistentes;
- Se optimiza el uso del tiempo de la Gerencia, que al implementar esta herramienta puede enfocarse a revisar aspectos de sus colaboradores como el desempeño individual de estos, optimizar los procesos en general e individualmente;
- Se asigna al mejor recurso, humano o *hardware*, para las tareas de mayor prioridad e importancia, los usuarios no toman tiempo en decidir qué tarea realizar ya que el sistema lo ha decidido en base a sus habilidades. No sólo es optimizar las tareas realizadas por el recurso humano. Cuando se habla de agilizar las tareas o procesos se incluye mover tanto a las personas como a los equipos, entre la sección, con otro departamento y hasta con otra compañía, todo con el fin de que se agilicen los procesos;
- Procesos paralelos o la ejecución simultánea de varios procesos se torna una tarea fácil de seguir, al llevarse en un sistema de flujos que en un flujo manual.

Otros beneficios serían:

- Re-asignación de puestos de trabajo redundantes
- Eliminar problemas de espacio
- Bajar el costo de los suministros
- Mejorar equipo obsoleto
- Integridad de la información
- Mejorar la productividad
- Dar rapidez al proceso
- Mejorar la moral de la gente
- Reducir costos
- Mejorar la imagen
- Proveer nuevos servicios

Uno de los problemas que se encuentra habitualmente en el desarrollo de aplicaciones para empresas, es que las tareas o procesos que se desarrollan en el entorno laboral de las mismas quedan inmersos en el código de la aplicación que resuelve la problemática de la empresa. Está claro que la gran mayoría de los usuarios no tienen conocimiento de estas tareas, las mismas están ocultas a sus ojos y se realizan automáticamente. El hecho de realizar cambios en dichas tareas o procesos resulta muy costoso, y es muy factible que dichos cambios redunden en realizar nuevamente la aplicación.

Una buena solución al problema anterior es separar los procedimientos y asociarlos a los flujos de trabajo realizados dentro de la empresa. Vemos entonces, que el Workflow se relaciona con la automatización de los procedimientos donde los documentos, la información o tareas son pasadas entre los participantes del sistema de acuerdo a un conjunto de reglas previamente establecidas. El fin de lo anterior es llegar a culminar una meta común impuesta por la empresa.

El workflow nos da las facilidades para modelar los procesos de empresa, permitiéndonos de esta forma hacer un análisis y diseño más profundo de los mismos. Vemos entonces al workflow no sólo como una tecnología que nos facilita el cambio, sino también, como una tecnología que nos da un marco de referencia para el análisis y diseño previo a la implantación de un sistema que implica la interacción de diversos procesos.

11.9 Tipos de Workflow

Workflow se define típicamente en tres amplias categorías que van desde los más complejos y estructurados hasta los que usan E-Mail (correo electrónico).

11.9.1 Workflow de Producción

El Workflow orientado a producción o transacciones es usado en aplicaciones tradicionales gobernadas por una serie de normas y procedimientos. Requieren personal para realizar tareas repetitivas en las cuales los documentos pueden requerir ser accedidos por pedido (días, meses o aún años después). Además, se requieren reglas para crear y mantener un registro de auditoría de cada documento. Ejemplos de este tipo incluyen líneas de crédito, reclamos, etc.

Los mayores fabricantes de soluciones de producción son:

Action Technologies, FileNet, IA, IBM, Keyfile, Recognition International, ViewStar, Wang Laboratoires y Xerox.

11.9.2 Workflow AD-HOC

El workflow orientado a proyectos o AD-HOC incluye un indefinido grupo de personas con fechas específicas para realizar tareas. Este tipo de workflow implica una gran cantidad de tiempo para su coordinación. Es típicamente de corta vida y desestructurado, variando mucho en su complejidad. Ejemplos de este tipo son: desarrollo de planes estratégicos, diseño de productos, evaluación de un producto, etc. muchos de los fabricantes de flujo de trabajo "ad hoc" son a su vez importantes firmas de correo electrónico, como **Banyan Systems, Lotus, Microsoft, Novell, Digital o Verimation**.

11.9.3 Workflow Administrativo

Esta categoría incluye tareas de rutina simples usando correo electrónico. El intercambio de información tiene lugar en forma electrónica.

11.10 Sistemas de Workflow

Los sistemas de workflow pueden ser descritos de acuerdo al tipo de proceso para el cual han sido diseñados. Así nosotros definimos tres tipos de *Workflow*

- Sistemas de *workflow* basados en imágenes: Son diseñados para permitir el flujo de papeles o documentos dentro de la empresa transformándolos en imágenes digitales. Estos fueron los primeros sistemas que recibieron gran aceptación.
- Sistemas de *workflow* basados en formas: Han sido diseñados para guiar las formas de forma inteligente. Las formas son guiadas hacia cualquier punto de la organización dependiendo de la información que se le sea ingresada. Además este tipo de sistemas puede avisar cuando una tarea tiene que realizarse. Este sistema provee mayor capacidad que los sistemas basados en imágenes.
- Sistemas de *workflow* basados en la coordinación: Están diseñados para facilitar el cumplimiento de las tareas dando un marco para la coordinación de las acciones. Este marco está dirigido para hacer frente al campo de lo concerniente a lo humano (procesos de negocios), en vez de optimización de procesos de material o de información. Estos sistemas tienen la capacidad de mejorar la productividad organizacional atendiendo las necesidades de los clientes en vez de automatizar procedimientos que no tienen relación alguna con la satisfacción del cliente.

11.11 Modelos de Arquitectura

Por su arquitectura, las aplicaciones de flujo de trabajo pueden ser clasificadas de acuerdo a los tres modelos cliente/servidor posibles: mensajería electrónica, bases de datos compartida y bases de datos distribuida. En todas ellas, la recepción de asignaciones de tareas y su realización tienen lugar en el cliente.

11.11.1 Modelo Basado en Correo Electrónico

Resulta muy apropiado para aplicaciones de envío de documentos a través de rutas, como autorización de gastos o aprobación de órdenes de compra. Prácticamente toda la funcionalidad reside y es ejecutada en el cliente y sólo las funciones de mensajería propiamente dichas recaen en el servidor.

Las rutas, roles y reglas son ejecutadas por uno o más procesos del cliente. Los datos, las reglas para procesar los documentos y el documento propiamente dicho residen en el depósito de entrada del usuario, mientras que las reglas del proceso pueden viajar con el documento o ser mantenidas en la aplicación de correo electrónico del cliente. Dada su estructura, este modelo soporta usuarios remotos y múltiples sistemas operativos de red y plataformas cliente, por lo que está especialmente indicado para soportar flujos de trabajo entre distintas organizaciones.

Su principal inconveniente, sin embargo, consiste en la complejidad de gestionar las reglas de flujo de trabajo. Como residen en múltiples localizaciones, cualquier cambio debe efectuarse en docenas o cientos de aplicaciones repartidas por toda la organización. Además suele resultar difícil localizar el estado o situación exacta de una porción de trabajo, salvo saber que se encuentra en el dispositivo de recepción de un usuario. Por otra parte, cuando se trabaja con documentos, estos sólo están disponibles para el usuario que lo recibe en ese momento, y son incapaces de recoger información de "metagestión" relacionada con procesos, información que ayuda a los usuarios a adoptar decisiones comerciales.

11.11.2 Modelo de Bases de Datos Compartidos

Todo el proceso se realiza en el cliente, aunque aquí los documentos no se envían de un lado a otro a través del sistema de correo, si bien éste puede utilizarse para avisar al usuario de que tiene una tarea de flujo de trabajo pendiente de realizar. Por el contrario, los documentos se almacenan en una base de datos compartida en la red, de forma que siempre están disponibles para, al menos, ser visualizados. Las reglas y los roles pueden ser almacenados separadamente de los documentos en la misma base de datos o en otra diferente, dentro del software cliente o conectados a los documentos propiamente dichos.

Uno de los puntos fuertes de este modelo es que permite disponer de acceso continuo a los documentos; además, ofrece mejores capacidades de gestión, por ejemplo, de reglas y de seguimiento de estatus. También estas bases de datos permiten almacenar información histórica sobre los flujos de trabajo, haciendo así posible la metagestión.

Su principal limitación consiste en que obliga a estar conectado a la base de datos si ésta no soporta funciones de replicación, aunque incluso en este caso no se asegura que se pueda acceder a la versión más reciente.

11.11.3 Modelo de Bases de Datos Distribuidas o Cliente/ Servidor

Ejecuta las reglas en el servidor. Este ejecuta procesos o agentes que determinan el paso siguiente a realizar en el proceso de flujo de trabajo y, además, puede monitorizar su estado y notificar a los participantes sobre cualquier tarea próxima o retrasada. Las aplicaciones de flujo basadas en servidor pueden integrarse también con otras fuentes de datos o aplicaciones.

Este modelo ofrece todos los beneficios de la arquitectura de bases de datos compartidas y se integra fácilmente con el sistema de correo electrónico. Su principal punto fuerte consiste en la capacidad de controlar la aplicación de flujo desde el servidor, lo que incluye la posibilidad de monitorizar y gestionar cada aplicación de flujo, así como una "metagestión" de todo el proceso comercial. Permite, además, añadir tareas de servidor que pongan en marcha, monitoricen y gestionen la aplicación de flujo en cuestión.

Sus limitaciones son similares a las de las bases de datos compartidas, aunque maneja mejor los clientes remotos. Generalmente, la aplicación servidor requiere funcionar sobre sistemas operativos multitarea, como OS/2 y Unix, distintos muy probablemente del sistema cliente.

11.12 Tendencias

Actualmente, están surgiendo nuevas ofertas de productos software dirigidas a usuarios finales y a los desarrolladores de aplicaciones de las empresas que suponen un nuevo estadio en la evolución del flujo de trabajo. Este nuevo software se caracteriza por tres tendencias claves: interfaces gráficas de usuario, estándares de bases de datos y entornos de desarrollo orientados a objeto. Al irse incorporando a los productos estas funciones y características, la tecnología de flujo de trabajo se convertirá en la base para la creación de sistemas de información a nivel de empresa.

Utilizando iconos, los desarrolladores de aplicaciones y los usuarios pueden crear diagramas de flujo y organigramas en pantalla para representar cualquier proceso de flujo de trabajo. Lo mejor de estas funciones gráficas es que permiten representar procesos complejos con un alto grado de abstracción. Los productos también deberán ofrecer un conjunto de interfaces de programación de aplicaciones (API) que permitan realizar llamadas a otras aplicaciones.

Cada flujo de trabajo consiste en tres elementos básicos: las reglas que gobiernan el proceso, la información que se está dirigiendo a través de las diversas rutas, y las medidas o patrones de proceso que se están utilizando para informar sobre el encaminamiento que se está realizando. En las primeras implementaciones de técnicas de flujo de trabajo, estas reglas y mediciones de procesos se almacenaban en un formato propietario. Sin embargo, en los tres últimos años están apareciendo productos que almacenan todos los parámetros en una base de datos estándar, como las de **Oracle**, **Sybase** e **Informix**.

Las bases de datos también pueden seguir el estado de los procesos de flujo de trabajo y mantener una ruta de auditoría histórica de cada transacción. Así, por ejemplo, permiten recoger estadísticas vitales sobre la eficiencia del proceso de flujo de trabajo (como los volúmenes y tiempos medios de proceso de transacciones) y almacenarlas con el fin de generar informes.

La integración de bases de datos es un imperativo ineludible para mantener la viabilidad a largo plazo de las aplicaciones de empresa. Y, a medida que vayan apareciendo los estándares, los repositorios de base de datos se convertirán en una forma de compartir reglas de proceso de flujo de trabajo entre diferentes aplicaciones.

La orientación a objeto es otra de las tendencias clave. Esta tecnología resulta ideal para las aplicaciones de flujo de trabajo, ya que combina objetos de información (documentos y datos) con las reglas de proceso y los parámetros que los gobiernan. Gracias a ella, el objeto se convierte en una combinación de información y proceso. Los procesos

comerciales pueden definirse más rápidamente con software orientado a objeto que con codificación tradicional, utilizando leyes de herencia, clases de objetos y encapsulación. Las bibliotecas de codificación ayudan a preservar los atributos de un objeto a través de todos los modelos de proceso. Además, pueden crearse superclases y subclases de objetos. Por ejemplo, las órdenes de compra pueden ser una subclase del modelo de proceso de aprovisionamiento, pero también una superclase para el proceso o tramitación de pedidos. La encapsulación enmascara toda esta complejidad dentro de un único objeto o ícono, que puede ser agrupado con cualquier combinación de otros objetos, para su incorporación a un nuevo flujo de trabajo.

Cuando se combina con un entorno de desarrollo gráfico, la tecnología de orientación a objeto ofrece un alto grado de integridad para las actividades de desarrollo de aplicaciones a nivel de empresa.

Asimismo, cada vez más, las técnicas de flujo de trabajo serán asimiladas por los sistemas operativos, integrándose así en los entornos de sobremesa. Esto será de una importancia crítica, porque aumentará el número de aplicaciones que ofrezcan el flujo de trabajo como una opción más. Y gracias a los emergentes estándares que están siendo desarrollados por la Workflow Management Coalition (WfMC), los usuarios podrán utilizar múltiples productos que funcionen conjuntamente y compartan objetos de flujo de trabajo, tal como pueden hacerlo hoy las bases de datos SQL.

Con el tiempo, la tecnología de flujo de trabajo se convertirá en la más difundida de todas las tecnologías de sobremesa y pasará a ser un ícono más en el entorno de trabajo del usuario final. Estos iconos o agentes realizarán las miles de tareas que están actualmente a cargo del empleado de oficina y decidirán automáticamente el paso de los documentos a otros empleados, de acuerdo con las reglas que representan los procesos que gobiernan a la empresa.

11.13 Resumen

Para desarrollar el Workflow, el primer paso consiste en poner en línea toda la información de negocios, el segundo paso es utilizar "**Workflow**" sistematizados para direccionar la información, como aprobaciones, modificaciones, a las personas apropiadas. Una vez definido el "**Workflow**", el sistema distribuirá la información a través de toda la organización, un programa monitor se encargará del control del progreso y estado de las transacciones.

En cada "**Workflow**", a través de las actividades previamente definidas, se efectuará el procesamiento por cada transacción desde su ingreso a la empresa hasta su culminación y/o archivo. Por ejemplo para Ventas el "**Workflow**" podrá tener como inicio la actividad de recepción del pedido del cliente y como fin la cancelación de la factura correspondiente, pasando por actividades intermedias como verificación y separación de la mercadería solicitada, la aprobación de crédito, despacho, descarga de inventarios, facturación, etc.

En cada actividad del "Workflow" se podrán definir las condiciones para su ejecución tales como responsable(s), requisitos previos, aprobaciones anteriores, etc. De esta manera, en todo momento, se dispondrá de información sobre la ubicación y el estado de la transacción dentro del "Workflow". Al ingresar o salir del sistema, o por intervalos de tiempo previamente definidos, se notificará automáticamente al (los) usuario(s) sobre las transacciones pendientes de su responsabilidad.

Mediante el "Workflow" se automatizan e integran el trámite documentario, los controles administrativos, aprobaciones y el procesamiento electrónico de los datos, los que tradicionalmente son ejecutados separadamente, demandando de procesos manuales adicionales que interfieren con la oportunidad y seguridad en el tratamiento de los datos. El Workflow "monitorea" cada transacción direccionando selectivamente el proceso a las personas competentes y notifica las interrupciones en la secuencia para su pronto reinicio o culminación. Esta capacidad permite, además, auditar completamente a una transacción, desde el evento inicial hasta su culminación (en ambos sentidos), con indicación de fechas, horas, personas que intervienen y tiempos insumidos, datos de gran relevancia para medir la eficiencia de toda la organización.

Manejando la información a través de límites organizacionales y actividades relacionadas en forma separada, como parte de un proceso coherente, las compañías pueden reducir los tiempos en estos ciclos críticos del negocio.

Estos nuevos sistemas de información deben manejar también la distribución de la información crítica, de manera que permita una acertada toma de decisiones de todos los miembros de la organización. Esta información compartida hace posible crear organizaciones orientadas a trabajar en equipos en lugar de modelos jerárquicos que en el pasado fueron necesarios debido al restringido acceso a la información.

First Order System Types

1. Introduction:

First order systems are, by definition, systems whose input-output relationship is a first order differential equation. A first order differential equation contains a first order derivative but no derivative higher than first order – the order of a differential equation is the order of the highest order derivative present in the equation.

First order systems contain a single energy storage element. In general, the order of the input-output differential equation will be the same as the number of independent energy storage elements in the system. Independent energy storage cannot be combined with other energy storage elements to form a single equivalent energy storage element. For example, we previously learned that two capacitors in parallel can be modeled as a single equivalent capacitor – therefore, a parallel combination of two capacitors forms a single independent energy storage element.

First order systems are an extremely important class of systems. Many practical systems are first order; for example, the mass-damper system and the mass heating system are both first order systems. Higher order systems can often be approximated as first order systems to a reasonable degree of accuracy if they have a dominant first order mode.

2. First Order System Model

The first order system has only one pole as shown

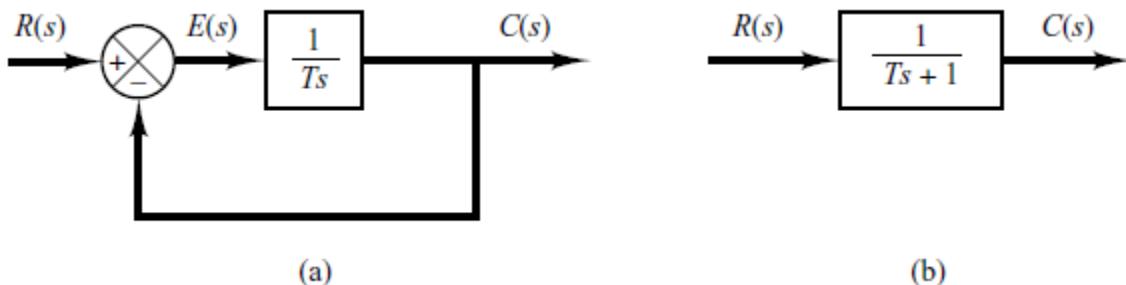


Figure 1: (a) Block Diagram of a first-order system; (b) Simplified block Diagram

$$\frac{C(s)}{R(s)} = K \frac{1}{Ts + 1} \quad (1)$$

- Where K is the DC Gain and T is the time constant of the system.
- Time Constant is a measure of how quickly a 1st order system response to a unit step input.
- DC gain of the system ration between the input signal and the steady state value of output.

Example 1:

For the first order system given below

$$G(s) = \frac{10}{3s + 1}$$

- DC gain (K) is equal 10
- Time Constant (T) is equal 3

Example 2:

For the first order system given below

$$G(s) = \frac{3}{s + 5} = \frac{\frac{3}{5}}{\frac{1}{5}s + 1}$$

- DC gain (K) is equal $\frac{3}{5}$
- Time Constant (T) is equal $\frac{1}{5}$

2.1 Impulse Response of 1st Order System

Consider the following 1st order system in figure 2

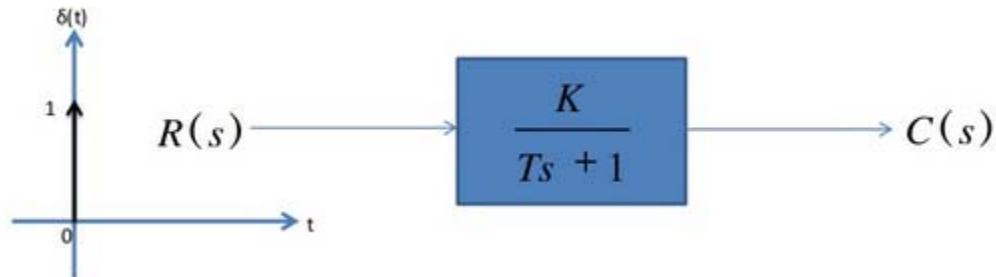


Figure 2: Impulse response of first order system

$$R(s) = \delta(s) = 1$$

$$C(s) = \frac{K}{Ts + 1}$$

In order to represent the response of the system in time domain we need to compute the inverse Laplace transform of the above equation, we have

$$c(t) = \frac{K}{T} e^{-\frac{t}{T}} \quad (2)$$

Example 3:

For the first order system given below

$$G(s) = \frac{3}{2s + 1}$$

The impulse response is shown in figure 3

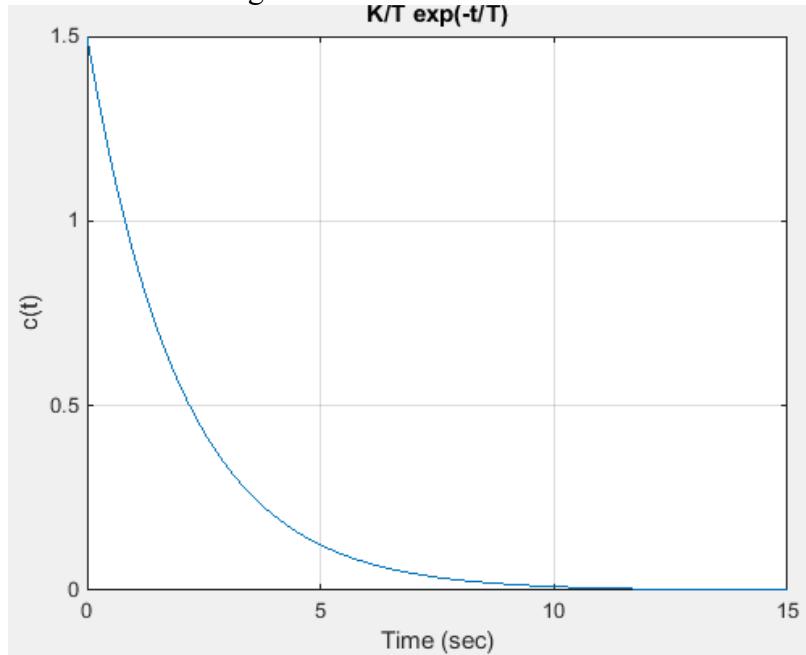


Figure 3: Impulse response of example 3

2.2 Ramp Response of 1st Order System

Consider the first order system in figure 1.

$$R(s) = \frac{1}{s^2}$$
$$C(s) = \frac{K}{Ts + 1} \frac{1}{s^2}$$

In order to represent the response of the system in time domain we need to compute the inverse Laplace transform of the above equation, we have

$$c(t) = Kt - T + Te^{-\frac{t}{T}} \quad (3)$$

Example 4:

For the first order system given below

$$G(s) = \frac{K}{Ts + 1}$$

- 1) The ramp response of system, if $K = 1$ and $T = 1$ is shown in figure 4
- 2) The ramp response of system, if $K = 1$ and $T = 3$ is shown in figure 5

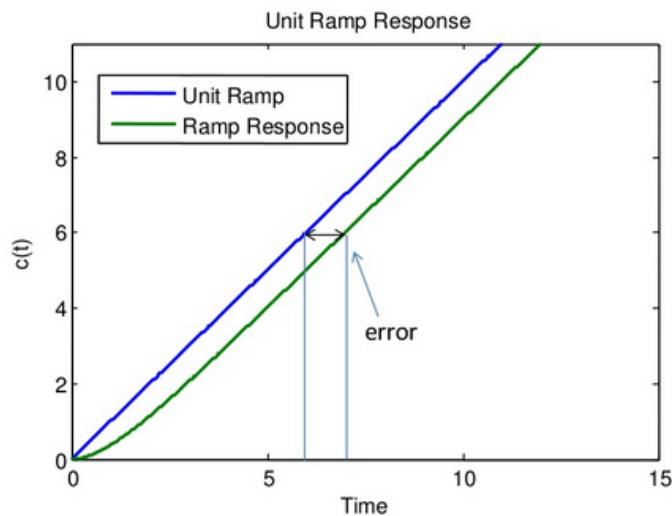


Figure 4: The ramp response for case 1, example 4

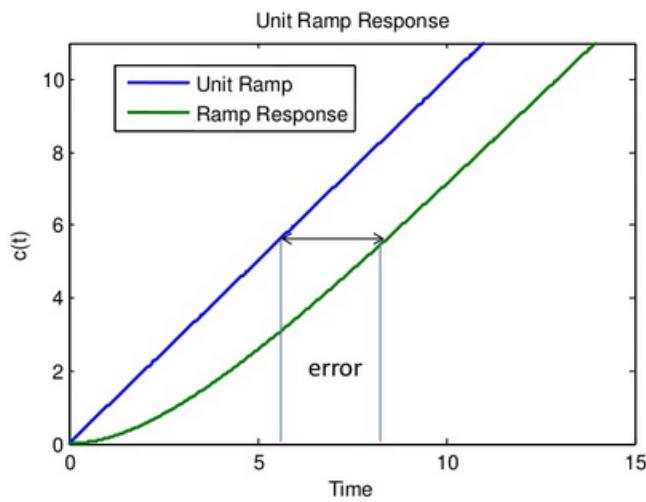


Figure 5: The ramp response for case 2, example 4

2.3 Step Response of 1st Order System

Consider the first order system in figure 1.

$$R(s) = \frac{1}{s}$$
$$C(s) = \frac{K}{Ts + 1} \frac{1}{s}$$

In order to represent the response of the system in time domain we need to compute the inverse Laplace transform of the above equation, we have

$$c(t) = Ku(t) - e^{-\frac{t}{T}} \quad (4)$$

- 1) Where $u(t) = 1$

$$c(t) = K - e^{-\frac{t}{T}} \quad (5)$$

- 2) Where $t = T$

$$c(t) = K - e^{-1} = 0.632K \quad (6)$$

- For example, assume $K = 10, T = 1.5s$

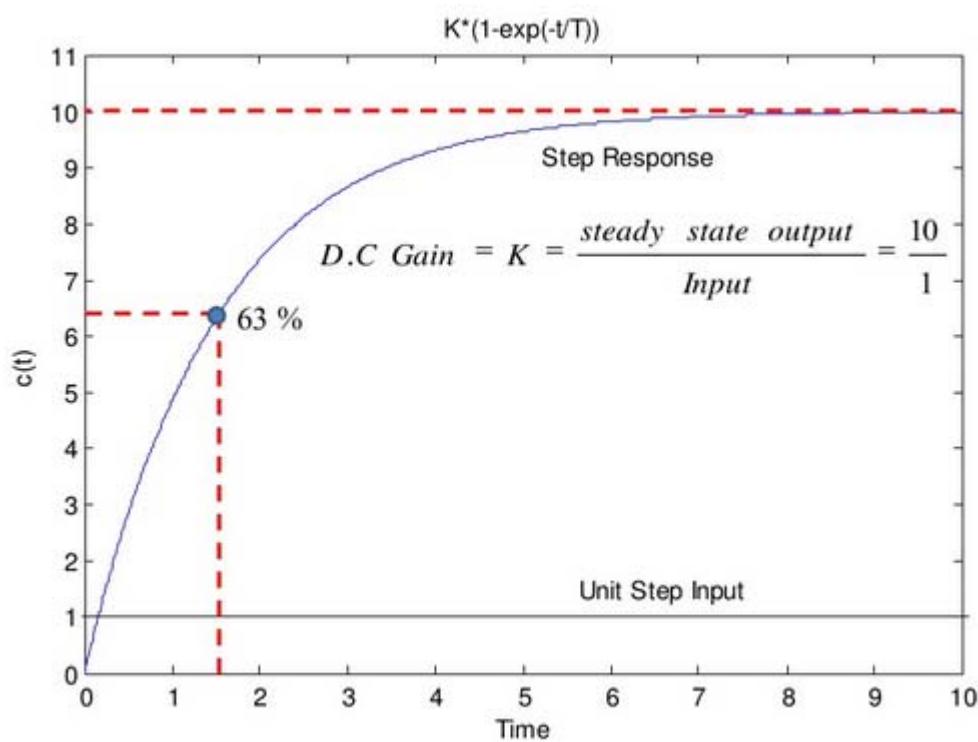


Figure 6: The step response specification of first order system

The step response of the first order system takes five time constants to reach its final value.

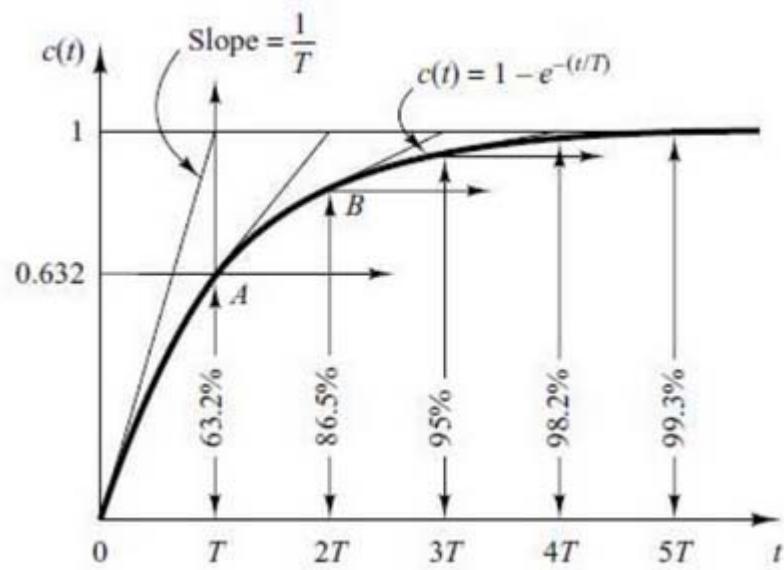


Figure 7: Step response

- $K = 10, T = 1, 3, 5, 7 \text{ s}$

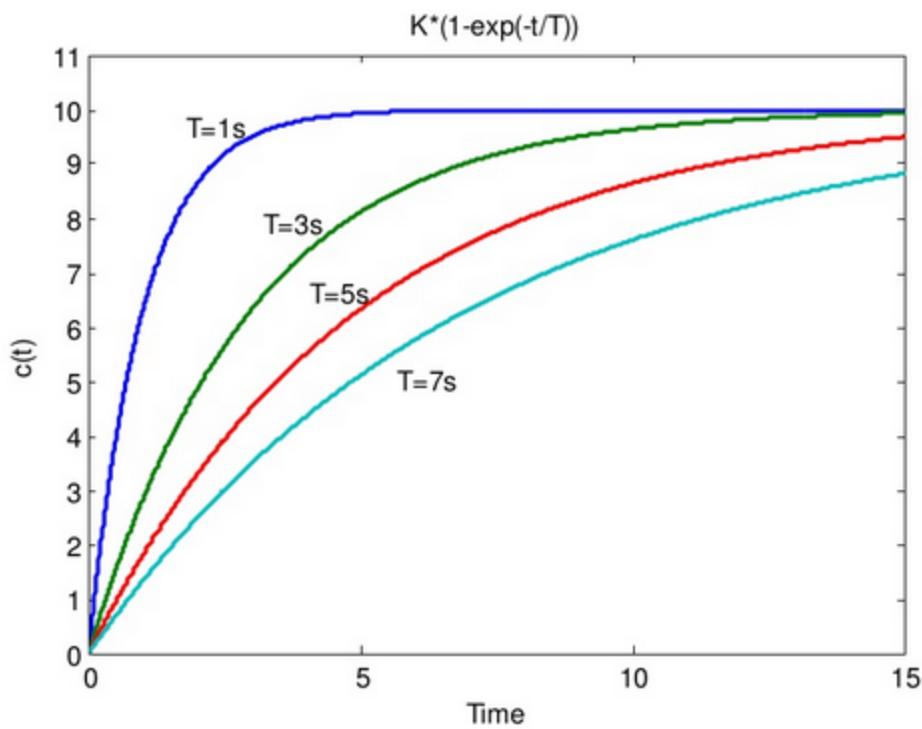


Figure 8: The step response at different value of T

- $K = 1, 3, 5, 10, T = 1 \text{ s}$

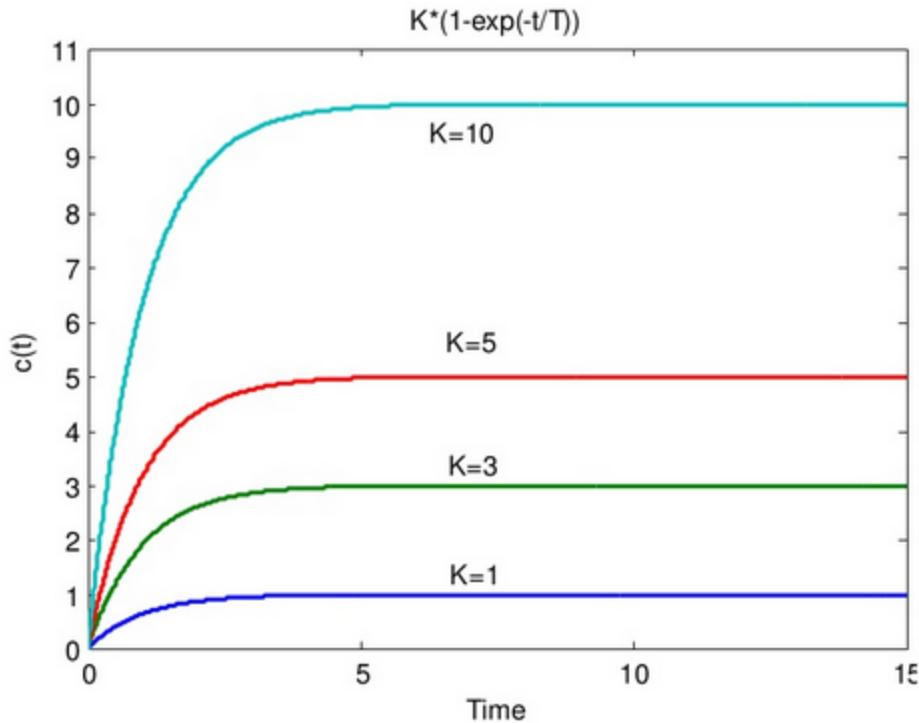


Figure 9: The step response at different value of K

3. First Order System with a Zero

The transfer function of the first order system with a zero can be represented using the form in equation (7)

$$G(s) = \frac{C(s)}{R(s)} = \frac{K(1 + \alpha s)}{Ts + 1} \quad (7)$$

- Zero of the system lie at $s = -1/\alpha$ and pole at $s = -1/T$
- Step response of the system would be:

$$C(s) = K \frac{1}{s} \frac{(1 + \alpha s)}{Ts + 1}$$

The Laplace inverse transfer is

$$c(t) = K + \frac{K}{T} (\alpha - T) e^{-\frac{t}{T}} \quad (8)$$

- **Case 1: $T > \alpha$**

The shape of the step response is approximately same (with offset added by zero)

Example 5:

Consider the first order system given by

$$\frac{C(s)}{R(s)} = \frac{10(1+2s)}{3s+1}$$

In this system:

1. $K = 10$
2. $T = 3$
3. $\alpha = 2$

$$c(t) = 10 + \frac{10}{3}(2 - 3)e^{-t/3}$$

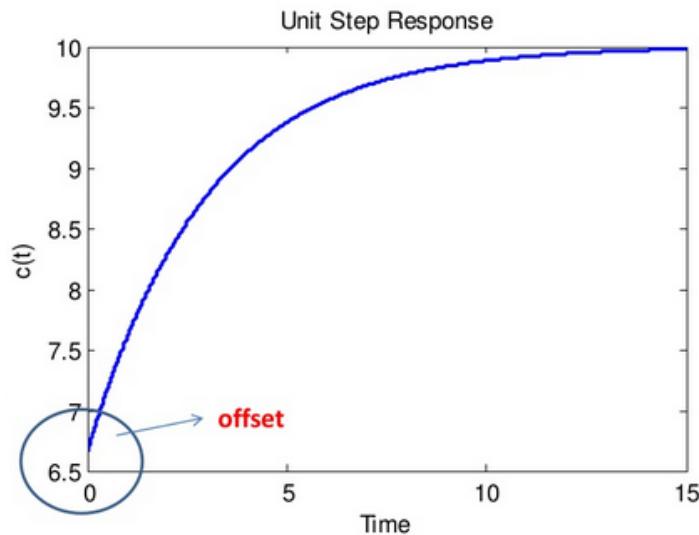


Figure 10: the step response of case 1 $T > \alpha$

$$\text{offset} = K + \frac{K}{T}(\alpha - T)$$

- **Case 2: $T < \alpha$**

Example 6:

Consider the first order system given by

$$\frac{C(s)}{R(s)} = \frac{10(1+2s)}{1.5s+1}$$

In this system:

1. $K = 10$
2. $T = 1.5$
3. $\alpha = 2$

$$c(t) = 10 + \frac{10}{1.5} (2 - 1)e^{-t/1.5}$$

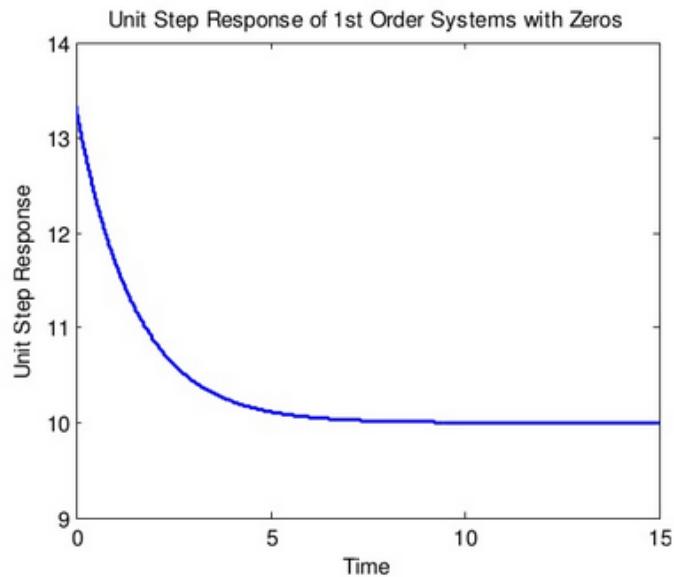


Figure 11: The step response for case 2 $T < \alpha$

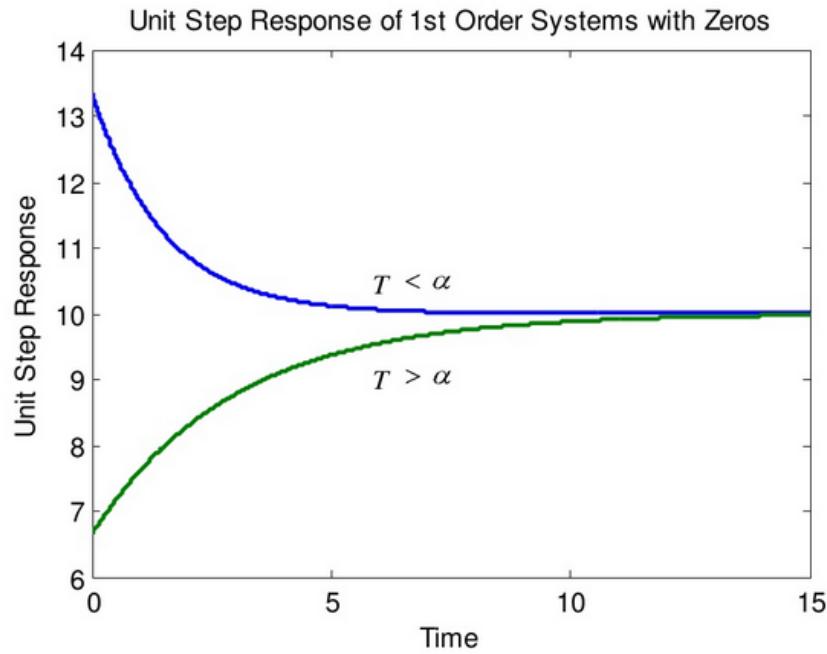


Figure 12: The step response of first order system with zero

4. First Order System with Delays

The first order system with delay time can have the following transfer function

$$\frac{C(s)}{R(s)} = \frac{K}{Ts + 1} e^{-st_d} \quad (9)$$

Where

t_d : is the delay time

The step response of this type of system is shown in figure 13

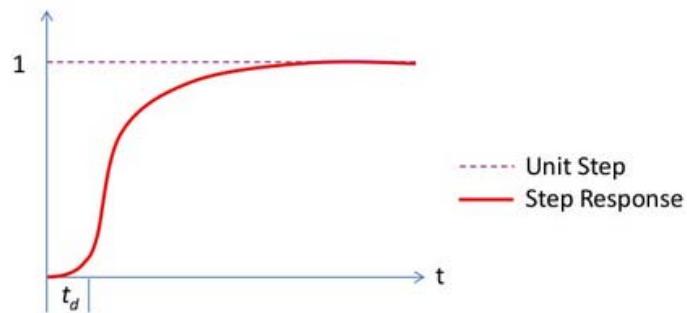


Figure 13: the step response of first order system with delay time

Example 6:

Consider the following first order system

$$G(s) = \frac{10}{s + 1} e^{-2s}$$

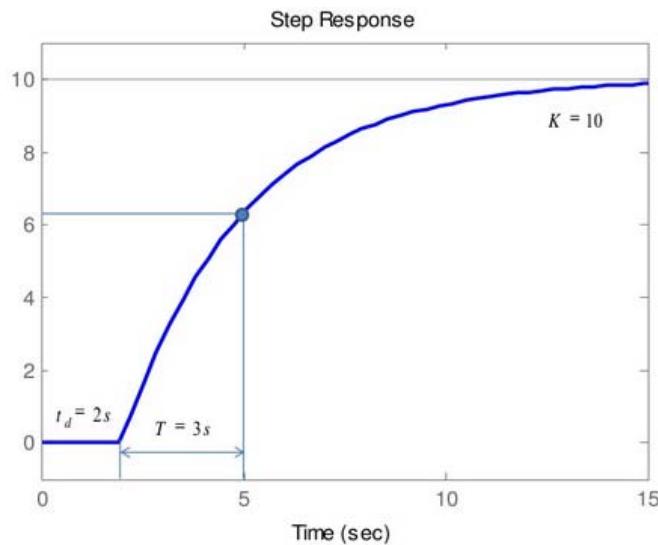


Figure 14: step response for system in Example 6

5. System Identification of Transfer Function of 1st Order Systems

The field of system identification uses statistical methods to build mathematical models of dynamical systems from measured data. System identification also includes the optimal design of experiments for efficiently generating informative data for fitting such models as well as model reduction.

A much more common approach is therefore to start from measurements of the behavior of the system and the external influences (inputs to the system) and try to determine a mathematical relation between them without going into the details of what is actually happening inside the system. This approach is called system identification. Two types of models are common in the field of system identification:

- 1) **Grey box Model:** although the peculiarities of what is going on inside the system are not entirely known, a certain model based on both insight into the system and experimental data is constructed. This model does however still have a number of unknown free parameters which can be estimated using system identification.

One example, uses the Monod saturation model for microbial growth. The model contains a simple hyperbolic relationship between substrate concentration and growth rate, but this can be justified by molecules binding to a substrate without going into detail on the types of molecules or types of binding. Grey box modeling is also known as semi-physical modeling.

- 2) **Black box Model:** No prior model is available. Most system identification algorithms are of this type.

In science, computing, and engineering, a black box is a device, system or object which can be viewed in terms of its inputs and outputs (or transfer characteristics), without any knowledge of its internal workings. Its implementation is "opaque" (black). Almost anything might be referred to as a black box: a transistor, algorithm, or the human brain.



Figure 15: Black Box

5.1 System Identification of First order system

- Often it is not possible or practical to obtain a system's transfer function analytically.
- Perhaps the system is closed, and the component parts are not easily identifiable.
- The system's step response can lead to a representation even though the inner construction is not known.
- With a step input, we can measure the time constant and the steady-state value, from which the transfer function can be calculated.
- If we can identify T and K from laboratory testing we can obtain the transfer function of the system.

Example 7:

Assume the unit step response given in figure 16;

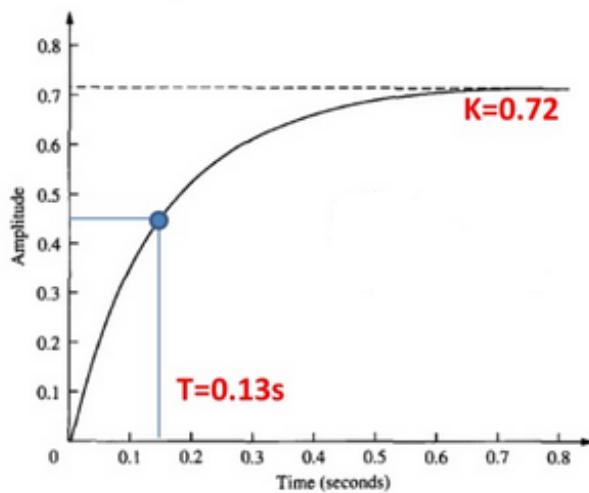


Figure 16: Unit step Response

- From the response, we can measure the **time constant T** , that is the time constant for the amplitude to reach 63% of its final value.

Since the final value is about 0.72 the time constant is evaluated for the curve reaches

$$0.63 \times 0.72 = 0.45, \text{ then } T = 0.13$$

- K is simply the steady state value.
- The transfer function is obtained as:

$$\frac{C(s)}{R(s)} = \frac{0.72}{0.13s + 1} = \frac{5.5}{s + 7.7}$$

6. Simulation of First order system using Simulink

In this section we study a open loop and closed loop system for case a first order system with delay and show the parameter of first order system.

Note: Maybe there more than blocks representation but we discuss and use the most model simulate the practical experiments as shown in the following example

Example 8:

Consider the first order system with delay given by the following transfer function

$$G(s) = \frac{2}{3s + 1} e^{-1s}$$

As shown in the previous section

$$K = 2 \text{ (DC gain)}$$

$$T = 3 \text{ (Time Constant)}$$

$$t_d = 1 \text{ (delay time)}$$

We can simulate this system in Simulink using the basic block diagrams (Transfer Fcn, gain , sum and Transport Delay)

- **Open Loop System**

We assume the system in unit step input.

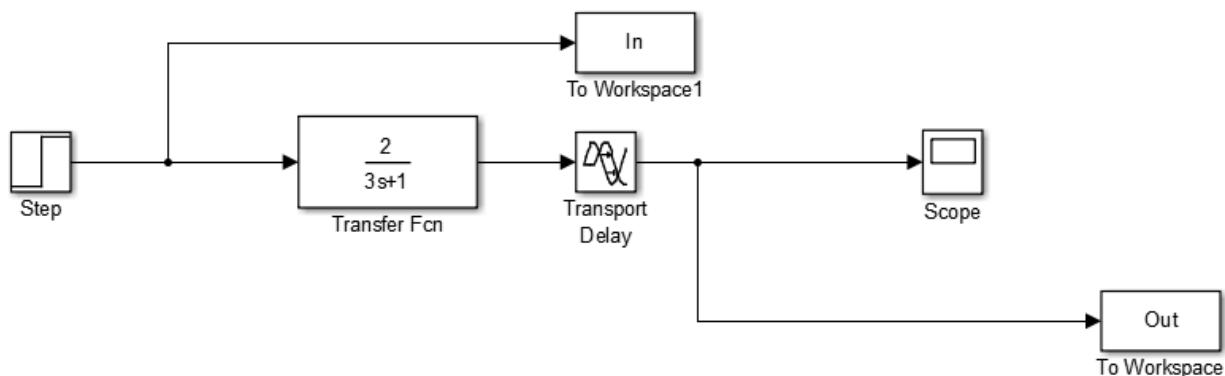


Figure 17: Open loop of system in example 8

The result of step response and the parameter is shown in figure 18

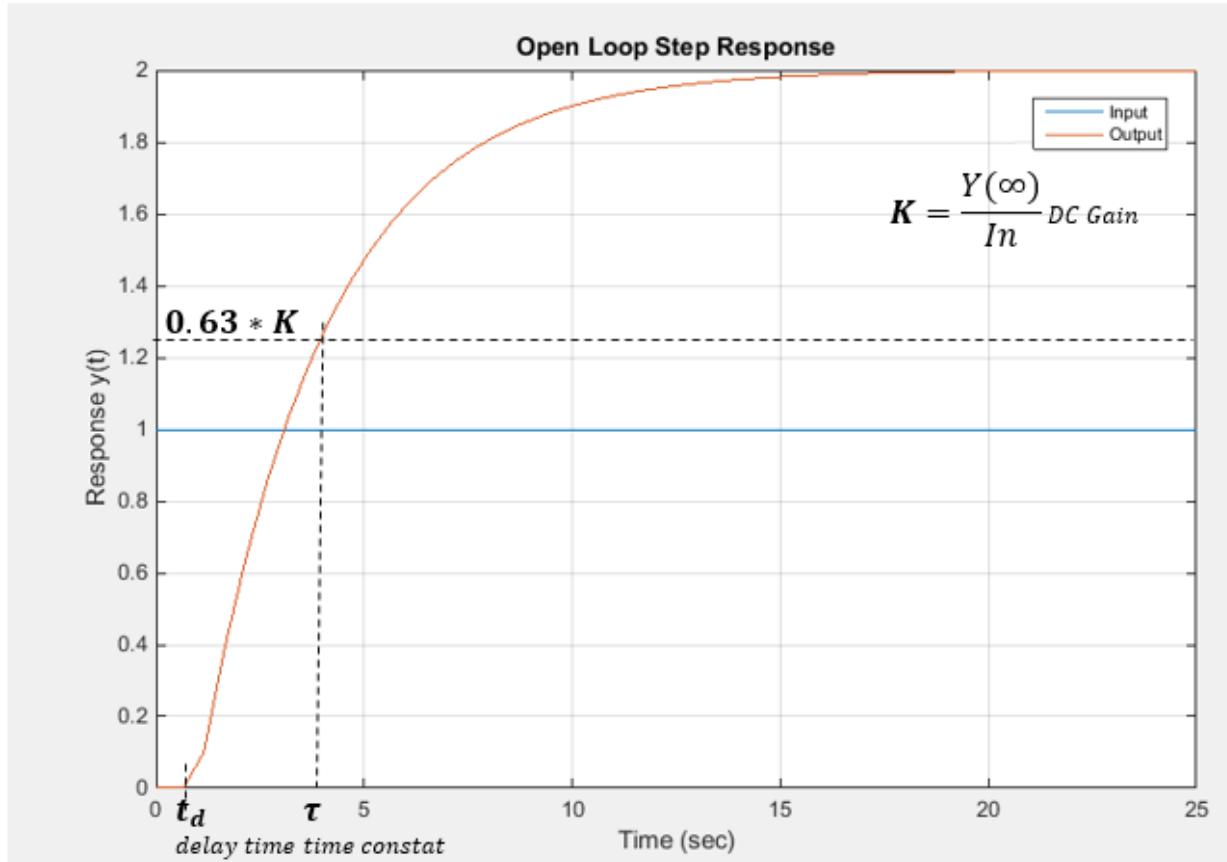


Figure 18: Open loop Response

So if the experimental data for first order system is known we can obtain the transfer function by determine the parameter (K, τ, t_d) as discussed in the system identification example.

- **Closed loop Response**

The block diagram representation of closed loop system is shown in the figure 19

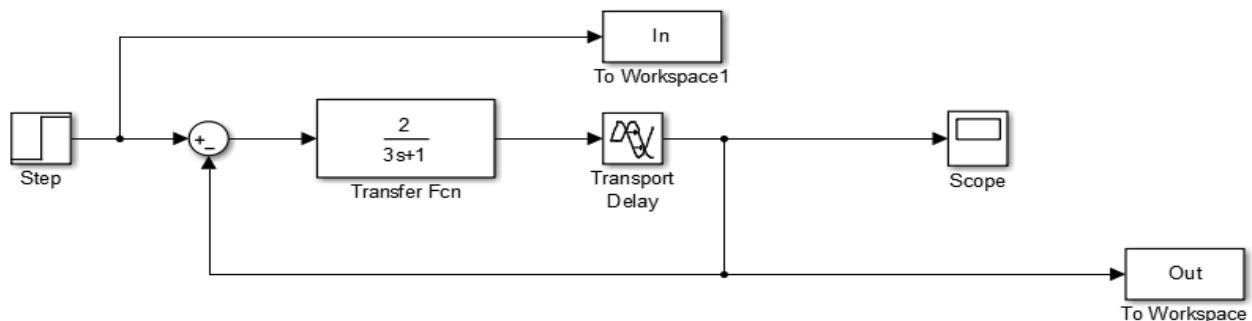


Figure 19: Closed Loop Bock Diagram

The step response is shown in figure 20

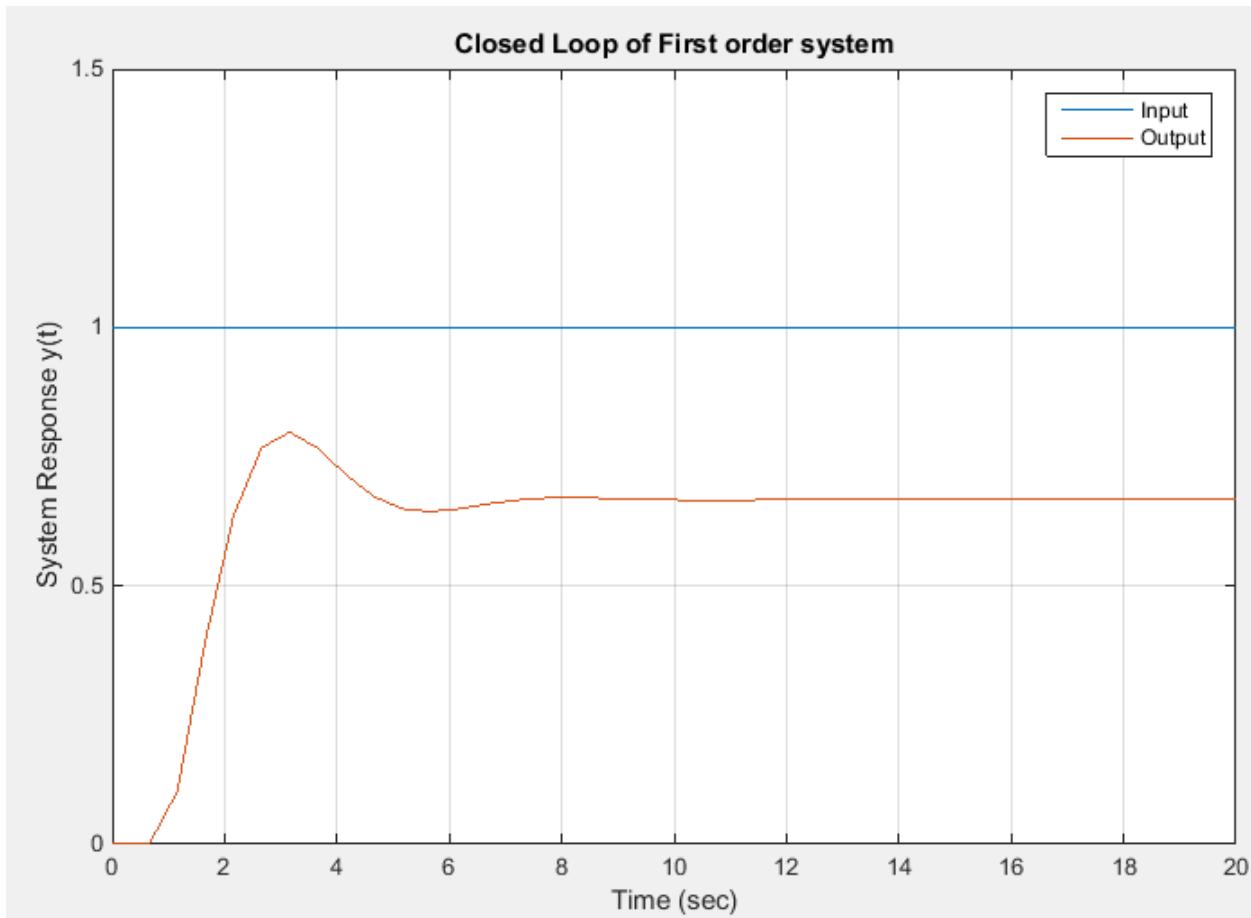


Figure 20: Step Response of closed loop

Depend on the two responses of system, we observe to determine the transfer function of the system we must get the parameter from the open loop characteristic because in this case we obtain the response and the behavior of the system without any modification or addition effect.