

ANNIVERSARY EDITION WITH FOUR NEW CHAPTERS



ESSAYS ON SOFTWARE ENGINEERING

THE MYTHICAL MAN-MONTH

FREDERICK P. BROOKS, JR.



Photo credit: ©Jerry Markatos

ACERCA DEL AUTOR

Frederick P. Brooks, Jr., es Profesor Kenan de Ciencias Computacionales de la Universidad de Carolina del Norte en Chapel Hill. Es mejor conocido como “el padre de la IBM Sytem/360,” se desempeñó como su director de desarrollo y más tarde como director de diseño de software del Operative System/360. Por este trabajo, en 1985, fue premiado junto a Bob Evans, y Erich Bloch con la Medalla Nacional de Tecnología. Previamente trabajó como arquitecto de las computadoras IBM Stretch y Harvest.

El Dr. Brooks fundó el Departamento de Ciencias Computacionales en Chapel Hill y lo presidió de 1964 a 1984. Ha servido en el Consejo Nacional de Ciencia y en el Consejo de Ciencia de la Defensa. Actualmente su actividad docente y de investigación se centra en la arquitectura de computadoras, las gráficas moleculares, y los ambientes virtuales.

El Mítico Hombre-Mes
Ensayos acerca de la
Ingeniería de Software
Edición de Aniversario

Frederick P. Brook, Jr.
Universidad de Carolina del Norte en Chapel Hill

Traducción:
Aldo Nuñez Tovar
anunez20@gmail.com

Dedicatoria de la edición de 1975

A dos personas que de forma especial enriquecieron mi estancia en IBM:

Thomas J. Watson, Jr.,

cuya profunda preocupación por la gente aún permea su empresa,

y

Bob O. Evans,

cuyo audaz liderazgo convirtió el trabajo en aventura.

Dedicatoria de la edición de 1995

A Nancy,

fue un regalo de dios.

Prefacio a la 20ma. Edición de Aniversario

Estoy sorprendido y satisfecho pues *El Mítico Hombre-Mes* continúa siendo popular después de 20 años. Se han imprimido más de 250,000 copias. Con frecuencia me preguntan qué opiniones y recomendaciones expuestas en 1975 todavía sostengo, y cuáles han cambiado, y cómo. De vez en cuando he abordado tales preguntas en conferencias, aunque desde hace tiempo he querido hacerlo por escrito.

Peter Gordon, ahora socio editorialista de Addison-Wesley, ha estado trabajando conmigo paciente y amablemente desde 1980. Me propuso preparar una Edición de Aniversario. Decidimos no revisar el original, sino reimprimirlo sin modificaciones (a excepción de algunas correcciones triviales) y añadirle reflexiones más actuales.

El capítulo 16 reimprime “No Existen Balas de Plata: Esencia y Accidentes de la Ingeniería de Software”, un artículo publicado en 1986 para la conferencia del IFIP que surgió a raíz de mi experiencia cuando presidí una investigación del Consejo de Ciencia de la Defensa acerca del software militar. Mis coautores de ese estudio, y nuestro secretario ejecutivo, Robert L. Patrick, fueron invaluableles al ponerme en contacto nuevamente con proyectos de software grandes del mundo real. El artículo fue reimprimido en 1987 en la revista *Computer magazine* de la IEEE, lo cual le dió una amplia difusión.

El artículo “No Existen Balas de Plata” demostró ser provocativo. Predijo que en una década no se observaría ninguna técnica de programación que trajera por sí misma una mejora de un orden de magnitud en la productividad

del software. Falta un año para la década; y mis predicciones parecen seguras. “No Existen Balas de Plata” ha estimulado mucho más el espíritu de discusión en la literatura que *El Mítico Hombre-Mes*. Por lo tanto, el Capítulo 17 habla acerca de algunas de las críticas publicadas y actualiza las opiniones expuestas en 1986.

Al preparar mi retrospectiva y actualización de *El Mítico Hombre-Mes*, me impresionó cuán pocas afirmaciones del libro han sido criticadas, demostradas, o refutadas por la actual investigación y experiencia de la ingeniería de software. Esto ahora me resulta útil para catalogar esas afirmaciones de forma cruda, desprovistas de hipótesis y datos de apoyo. Con la esperanza de que estas francas declaraciones inviten a debatir y a los hechos a demostrar, refutar, actualizar, o mejorar dichas proposiciones, las he incluido en el Capítulo 18 en forma de resumen.

El Capítulo 19 es en sí el ensayo de actualización. Se advierte al lector que estas nuevas opiniones no están ni de cerca tan bien respaldadas por la experiencia en las trincheras como en el libro original. He estado trabajando en una universidad, no en una industria, y en proyectos de pequeña escala, no en grandes. Desde 1986, solo he enseñado ingeniería de software, no he realizado investigación en este campo en absoluto. Mi investigación más bien se ha centrado en los ambientes virtuales y sus aplicaciones.

En la preparación de esta retrospectiva, he buscado las opiniones actualizadas de amigos que de hecho están trabajando en la ingeniería de software. Por la maravillosa buena voluntad de compartir sus puntos de vista, a comentar concienzudamente los borradores, y a reeducarme, estoy en deuda con Barry Boehm, Ken Brooks, Dick Case, James Coggins, Tom DeMarco, Jim McCarthy, David Parnas, Earl Wheeler, y Edward Yourdon. Fay Ward se ha encargado de forma magnífica de la producción técnica de los nuevos capítulos.

Agradezco a Gordon Bell, Bruce Buchanan, Rick Hayes-Roth, mis colegas en la Fuerza de Tarea del Comité de Ciencia de la Defensa acerca del Software Militar, y, muy especialmente, a David Parnas por sus profundas y estimulantes ideas, y a Rebekah Bierly por la producción técnica del artículo imprimido aquí como el Capítulo 16. Analizar el problema del software en categorías de *esencia* y *accidente* fue inspirado por Nancy Greenwood Brooks, que utilizó dicho análisis en un artículo acerca de la pedagogía del violín Suzuki.

En el prefacio de la edición de 1975 los derechos de la casa editorial Addison Wesley no me permitieron reconocer los roles claves jugados por su personal. Debo citar especialmente las contribuciones de dos personas: Norman Stanton, Editor Ejecutivo en ese entonces, y Herber Boes, Director de Arte en aquella época. Boes diseñó el estilo elegante, fue especialmente mencionado por un revisor: “amplios márgenes, [y] un uso imaginativo de la tipografía y la plantilla.” Y más importante aún, él también me hizo la recomendación crucial de que cada capítulo tuviera un dibujo inicial. (En ese entonces solo tenía el Pozo de Brea y la Catedral de Reims.) Encontrar los dibujos me ocasionó un año extra de trabajo, pero estoy eternamente agradecido por el consejo.

Soli Deo gloria— Gloria solo a dios

*Chapel Hill, N.C.,
Marzo de 1995*

F. P. B., Jr.

Prefacio a la Primera Edición

En muchos sentidos, la gestión de un proyecto grande de programación de sistemas es como gestionar cualquier otra gran empresa –en más formas de las que la mayoría de los programadores creen. Pero en muchos otros sentidos es diferente–en más formas de las que la mayoría de los gestores esperan.

El conocimiento del área se ha estado acumulando. Se han llevado a cabo varios congresos, sesiones en congresos de la AFIPS, y publicado algunos libros, y artículos. Pero de ninguna manera aún en forma de un enfoque sistemático de libro de texto. Sin embargo, me parece apropiado ofrecer este pequeño libro, que refleja esencialmente un punto de vista personal.

Aunque originalmente me desarrollé en el lado de la programación, estuve principalmente involucrado en la arquitectura de hardware durante los años (1956-1963), en los que se desarrolló el programas de control autónomo y el compilador de lenguajes de alto nivel. Cuando en 1964 llegué a ser director del Operative System/360, encontré el mundo de la programación bastante cambiado debido al progreso de los últimos años.

Gestionar el desarrollo del OS/360 fue una experiencia muy educativa, aunque también algo muy frustrante. El equipo, incluyendo a F. M. Trapnell que me sucedió como director, tiene mucho de que enorgullecerse. El sistema contiene muchas virtudes en su diseño y ejecución, y ha tenido éxito en lograr un amplio uso. Ciertas ideas, las más evidentes son las entradas y salidas independientes del dispositivo y la gestión de bibliotecas externas, fueron innovaciones técnicas y que hoy en día son copiadas ampliamente. Actualmente

es bastante confiable, razonablemente eficiente, y muy versátil.

Sin embargo, no puede decirse que el esfuerzo fue totalmente exitoso. Cualquier usuario del OS/360 se percata rápidamente de cuánto mejor debería ser. Los defectos en el diseño y la ejecución permearon especialmente al programa de control, a diferencia de los compiladores. La mayoría de estos defectos datan del periodo de diseño de 1964-65, por tanto son de mi responsabilidad. Además, el producto estaba retrasado, tomó más memoria de lo planificado, los costos superaron en mucho lo estimado, y no funcionó muy bien hasta varias entregas posteriores a la primera.

Luego de dejar IBM en 1965 para venir a Chapel Hill, como originalmente acordamos cuando asumí el cargo del OS/360, empecé a analizar la experiencia del OS/360 para ver qué lecciones técnicas y de gestión había que aprender. En particular, quería explicar las amplias diferencias en las experiencias de gestión encontradas en el desarrollo del hardware en la System/360 y el desarrollo del software del OS/360. Este libro es una respuesta tardía al cuestionario de Tom Watson de por qué la programación es difícil de gestionar.

En esta búsqueda me he beneficiado de largas conversaciones con R.P. Case, director adjunto de 1964-65, y con F. M. Trapnell, director de 1965-68. He comparado resultados con otros directores de proyectos grandes de programación, incluyendo a F.J. Corbato del M.I.T., a Jhon Harr y V. Vyssotsky de Bell Telephone Laboratories, a Charles Portman de International Computer Limited, a A. P. Ershov del Laboratorio de Computación de la División Siberiana, de la Academia de Ciencias de la U.R.S.S., y a A. M. Pietrasanta de IBM.

Mis propias conclusiones están plasmadas en los siguientes ensayos, y están dirigidas a los programadores profesionales, a los gestores profesionales, y especialmente a los gestores profesionales de la programación.

Aunque escrito como ensayos separables, existe un tema principal contenido especialmente en los Capítulos del 2 al 7. En resumen, creo que los grandes proyectos de programación sufren de problemas de gestión de diferente tipo que los pequeños, a causa de la división del trabajo. Creo que la necesidad urgente es la preservación de la integridad conceptual del producto mismo. Estos capítulos exploran las dificultades por alcanzar esta unidad y los métodos para hacerlo. Los últimos capítulos exploran otros aspectos de la gestión de la ingeniería de software.

La literatura en este campo no es abundante, y está ampliamente dis-

persa. Por lo tanto, he tratado de dar referencias que por un lado aclaren puntos particulares y por otro guíen al lector interesado a otras obras útiles. Muchos amigos han leído el manuscrito, y algunos de ellos han preparado extensos comentarios provechosos; donde estos parecían valiosos pero no se ajustaban al flujo del texto, los incluía en las notas.

Debido a que este es un libro de ensayos y no uno de texto, todas las referencias y notas han sido desterradas al final del libro y se exhorta al lector a pasarlas por alto en una primera lectura.

Estoy profundamente en deuda con la Srita. Sara Elizabeth Moore, con el Sr. David Wagner, y con la Sra. Rebecca Burris por su ayuda en la preparación de este manuscrito, y con el Profesor Joseph C. Sloane por su consejo acerca de las ilustraciones.

Chapel Hill, N.C.
Octubre de 1974

F.P.B., Jr

Contenido

	<i>Prefacio a la 20ma. Edición de Aniversario</i>	vii
	<i>Prefacio a la Primera Edición</i>	xi
Capítulo 1	<i>El Pozo de Brea</i>	3
Capítulo 2	<i>El Mítico Hombre-Mes</i>	13
Capítulo 3	<i>El Equipo Quirúrgico</i>	29
Capítulo 4	<i>Democracia, Aristocracia y Diseño de Sistemas</i>	41
Capítulo 5	<i>El efecto del Segundo-Sistema</i>	53
Capítulo 6	<i>Pasar la Palabra</i>	61
Capítulo 7	<i>¿Por Qué Fracasó la Torre de Babel?</i>	73
Capítulo 8	<i>Predecir el Batazo</i>	87
Capítulo 9	<i>Diez Libras en un Saco de Cinco Libras</i>	97

1

El Pozo de Brea



1

El Pozo de Brea

*Enn schip op het strand is een baken in zee.
[Un barco en la playa es un faro para la mar.]*

PROVERBIO HOLANÉS

C.R. Knight, Mural de La Brea Tar Pits

El Museo de George C. Page de La Brea Discoveries

Museo de Historia Natural del Condado de Los Ángeles

Ninguna escena de la prehistoria es tan intensa como aquella en donde grandes bestias luchan mortalmente en los pozos de brea. Uno se imagina a dinosaurios, mamuts, y tigres dientes de sable luchando en contra de la adhesión de la brea. Pero cuanto más ferozmente luchan, más los enreda ésta, y ninguna bestia es tan fuerte o tan hábil para que finalmente no se hunda.

A lo largo de la década pasada la programación de sistemas grandes ha estado en tal pozo de brea, donde muchas bestias grandes y poderosas han sucumbido violentamente. A pesar de que la mayoría han emergido con sistemas que funcionan, unas pocas han cumplido con los objetivos, calendarios, y presupuestos. Grandes y pequeños, masivos o ligeros, equipo tras equipo se han enredado en la brea. Se puede apartar cualquier garra, pues nada parece causar el problema. Pero la acumulación de factores simultáneos e interactivos produce un movimiento cada vez más lento. Todos parecen estar sorprendidos por la persistencia del problema, y ha sido difícil discernir su naturaleza. Pero para solucionar el problema debemos tratar de entenderlo.

Por lo tanto, comencemos identificando el oficio de la programación de sistemas sus satisfacciones y sus infortunios intrínsecos.

El Producto de los Sistemas de Programación

De vez en cuando leemos relatos en las noticias acerca de como dos programadores en un remodelado garaje han construido un programa importante que supera los mejores logros de los equipos grandes. Y todo programador está dispuesto a creer tales historias, pues sabe que podría construir *cualquier* programa mucho más rápido que las 1000 declaraciones/año reportadas por los equipos industriales.

¿Por qué entonces no se han reemplazado todos los equipos industriales por dedicados dúos de garaje? Echemos un vistazo a *eso* que se está produciendo.

En la parte superior izquierda de la Fig. 1.1 está un *programa*. Está completamente terminado, listo para ser ejecutado por el autor en el sistema en el que fue desarrollado. *Ese* es el objeto que comúnmente se produce en los garajes, y es el que utiliza el programador para evaluar la productividad.



Fig. 1.1 Evolución del producto de los sistemas de programación

Hay dos formas en que un programa puede convertirse en un objeto más útil, pero más costoso. Estas dos formas están representadas por las fronteras del dibujo.

Si nos movemos hacia abajo y cruzamos la frontera horizontal, un programa llega a ser un *producto de programación*. Es decir, un programa que puede ser ejecutado, probado, reparado, y extendido por cualquiera. Es utilizable en muchos ambientes operativos y para muchos conjuntos de datos. Para convertirse en un producto de programación que pueda usarse de forma general, un programa debe estar escrito en un estilo generalizado. En especial, la gama y la forma de las entradas deben ser tan generalizadas como razonablemente lo permita el algoritmo básico. Luego, el programa debe ser

probado minuciosamente, de modo que podamos contar con él. Esto significa que se debe preparar, ejecutar y grabar un considerable banco de casos de prueba, que explore la gama de entradas y pruebe sus límites. Finalmente, el ascenso de un programa a un producto de programación requiere de una documentación minuciosa, tal que cualquiera pueda usarlo, corregirlo, y extenderlo. De modo empírico, calculo que un producto de programación cuesta al menos tres veces más que un programa depurado con la misma funcionalidad.

Si nos movemos a través de la frontera vertical, un programa se convierte en un componente dentro de un *sistema de programación*. Esto es, en una colección de programas interactivos, coordinados en funcionamiento y disciplinados en formato, tal que todo el ensamblaje constituya un instrumento completo para realizar tareas grandes. Un programa, para convertirse en un componente de un sistema de programación, debe escribirse de tal manera que toda entrada y toda salida se ajusten en sintaxis y semántica con interfaces muy bien definidas. El programa también debe diseñarse de tal manera que utilice un presupuesto determinado de recursos – espacio de memoria, dispositivos de entrada y salida, tiempo de cómputo. Por último, el programa debe probarse junto con otros componentes del sistema, en todas las combinaciones esperadas. Esta prueba debe ser amplia, debido al número de casos que crece de forma combinatoria. Todo esto consume tiempo, debido a los sutiles errores que surgen de las interacciones inesperadas de los componentes depurados. Un componente de un sistema de programación cuesta tres veces más que un programa autónomo con la misma funcionalidad. El costo puede ser mayor si el sistema tiene muchos componentes.

En la esquina inferior derecha de la Fig 1.1 se sitúa el *producto de los sistemas de programación*. Éste difiere de los programas simples en todas las formas de arriba. Cuesta nueve veces más. Pero realmente es un objeto útil, el producto planeado por el trabajo de gran parte de la programación de sistemas.

Las Satisfacciones del Oficio

¿Por qué la programación es divertida? ¿Qué satisfacciones pueden esperar sus practicantes como recompensa?

Primero, la pura satisfacción de hacer cosas. Así como los niños se deleitan haciendo sus pasteles de lodo, así los adultos se divierten construyendo cosas, especialmente cosas de sus propios diseños. Pienso que esta satisfacción debe ser una imagen de la satisfacción de dios de construir cosas, una satisfacción demostrada en la individualidad y originalidad de cada hoja y cada copo de nieve.

Segundo, es la satisfacción de hacer cosas que sean útiles para otras personas. En lo más profundo, queremos que otros usen nuestro trabajo y lo encuentren útil. En este sentido la programación de sistemas no es esencialmente diferente al primer recipiente de arcilla del niño para los lápices “de la oficina de papá”.

Tercero, es la fascinación de crear objetos complejos parecidos a rompecabezas, de entrelazar piezas móviles y observarlos trabajar en ciclos ingeniosos, ejecutando desde el inicio los resultados de sus fundamentos incorporados. Una computadora programada tiene toda la fascinación del juego de pinball o del mecanismo de la máquina de discos, llevada al extremo.

Cuarto, es la satisfacción del continuo aprendizaje, el cual brota de la naturaleza no repetitiva del trabajo. De una forma u otra el problema es siempre nuevo, y quien lo resuelve aprende algo: algunas veces algo práctico, otras algo teórico, y a veces ambos.

Finalmente, existe la satisfacción de trabajar en un medio tan manipulable. El programador, como el poeta, trabaja sólo ligeramente alejado del pensamiento puro. Construye sus castillos en el aire, a partir del aire, crea a través del esfuerzo de la imaginación. Pocos medios de la creación son tan flexibles, tan fáciles de pulir y reutilizar, tan fácilmente capaces de llevar a cabo estupendas estructuras conceptuales. (Como veremos después, esta gran manipulabilidad tiene sus propios problemas.)

Sin embargo, la construcción de programas a diferencia de las palabras del poeta, es real en el sentido que se mueve y trabaja, y produce salidas visibles separadas de la construcción misma. Imprime resultados, traza dibujos, produce sonidos, mueve brazos. La magia del mito y la leyenda se hace realidad en nuestro tiempo. Uno teclea el conjuro correcto, y la pantalla cobra

vida, mostrando cosas que jamás existieron ni podrían existir.

Por lo tanto, la programación es divertida porque gratifica el anhelo creativo inmerso en lo profundo de nosotros y alegra los sentimientos que tenemos en común con todas las personas.

Los Infortunios del Oficio

Sin embargo, no todo es satisfacción y conocer los infortunios intrínsecos nos ayudarán a lidiar con ellos cuando surjan.

En primer lugar, se debe actuar perfectamente. La computadora también se parece a la magia de la leyenda en este aspecto. Si un personaje o una pausa del conjuro no está estrictamente en la forma apropiada, la magia no funciona. Los seres humanos no estamos acostumbrados a ser perfectos, y pocas áreas de la actividad humana lo exigen. Pienso que ajustarse al requisito de perfección es la parte más difícil de aprender a programar.¹

Luego, otras personas nos asignan los objetivos, nos proveen los recursos, y nos facilitan la información. Uno rara vez controla las circunstancias de su trabajo, o incluso su objetivo. En términos de gestión, nuestra autoridad no es suficiente para toda nuestra responsabilidad. Sin embargo, parece que en todos los campos, donde se lleva a cabo el trabajo productivo no se tiene una autoridad formal proporcional a la responsabilidad. En la práctica, la autoridad real (como opuesta a la formal) se adquiere del mismo impulso del cumplimiento.

La dependencia de otros tiene una particularidad especialmente desagradable para el programador de sistemas. Pues éste depende de los programas de otras personas. Y estos están generalmente mal diseñados, mal implementados, liberados de forma incompleta (sin código fuente ni casos de prueba), y mal documentados. De tal manera que se debe invertir horas estudiando y corrigiendo cosas que en un mundo ideal deberían estar completas, disponibles, y ser utilizables.

El siguiente infortunio es que diseñar conceptos estupendos es divertido; encontrar una plaga de pequeños errores es solo trabajo. Acompañada de cualquier actividad creativa vienen monótonas horas de tedio, de meticuloso trabajo, y la programación no es la excepción.

Luego, uno descubre que la depuración tiene una convergencia lineal o peor, donde uno esperaba, por algún motivo, un tipo de convergencia cuadrá-

tica hacia el final. Así, las pruebas se hacen interminables, encontrar los últimos errores difíciles toma más tiempo que encontrar los primeros.

Y para empeorar las cosas, el último infortunio es que el producto en el que se ha trabajado durante tanto tiempo parece ser obsoleto al finalizar (o antes). Los colegas y competidores ya están en la intensa búsqueda de nuevas y mejores ideas. La actividad del pensamiento infantil ya no solo está concebida sino calendarizada.

Esto siempre parece peor de lo que en realidad es. El producto nuevo y mejorado generalmente no está *disponible* cuando uno termina el propio; de esto solo se habla. Y esto también requerirá meses de desarrollo. Jamás equiparemos un tigre de verdad con uno de papel, salvo que queramos usarlo de verdad. Por lo tanto, las ventajas de la realidad se satisfacen por sí mismas.

Por supuesto que la base tecnológica sobre la cual uno trabaja avanza *siempre*. Tan pronto como se congela el diseño, llega a ser obsoleto en términos de sus conceptos. Pero la implementación de productos reales exige coordinación y cuantificación. La obsolescencia de una implementación se debe medir contra otras implementaciones existentes, no contra conceptos no realizados. El reto y la misión son encontrar soluciones reales a problemas reales en base a calendarios reales de acuerdo a los recursos disponibles.

Pues esto es la programación, por un lado el pozo de brea, en el cual muchos esfuerzos han sucumbido, y por otro, la actividad creativa con sus propias satisfacciones e infortunios. Para muchos, las satisfacciones superan con creces los infortunios, y para ellos el resto de este libro intentará tender algunos puentes a través de la brea.

2

El Mítico Hombre-Mes

Restaurant Antoine

Fondé En 1840

AVIS AU PUBLIC

Faire de la bonne cuisine demande un certain temps. Si on vous fait attendre, c'est pour mieux vous servir, et vous plaire.

ENTREES (SUITE)

Côtelettes d'agneau grillées 2.50	Entrecôte marchand de vin 4.00
Côtelettes d'agneau aux champignons frais 2.75	Côtelettes d'agneau maison d'or 2.00
Filet de boeuf aux champignons frais 4.75	Côtelettes d'agneau à la parisienne
Ris de veau à la financière 2.00	Fois de volaille à la brochette 1.50
Filet de boeuf nature 3.75	Tournedos nature 2.75
Tournedos Médicis 3.25	Filet de boeuf à la hawaïenne 4.00
Pigeonneaux sauce paradis 3.50	Tournedos à la hawaïenne 3.25
Tournedos sauce béarnaise 3.25	Tournedos marchand de vin 3.25
Entrecôte minute 2.75	Pigeonneaux grillés 3.00
Filet de boeuf béarnaise 4.00	Entrecôte nature 3.75
Tripes à la mode de Caen (commander d'avance) 2.00	Châteaubriand (30 minutes)

LÉGUMES

Epinards sauce crème .60	Chou-fleur au gratin .60
Broccoli sauce hollandaise .80	Asperges fraîches au beurre .90
Pommes de terre au gratin .60	Carottes à la crème .60
Haricots verts au beurre .60	Pommes de terre soufflées
Petits pois à la française .75	

SALADES

Salade Antoine .60	Fonds d'artichauts Mayard
Salade Mirabeau .75	Salade de laitue aux oeufs .60
Salade laitue au roquefort .80	Tomate frappée à la Jules César .60
Salade de laitue aux tomates .60	Salade de coeur de palmier 1.00
Salade de légumes .60	Salade aux pointes d'asperges .60
Salade d'anchois 1.00	Avocat à la vinaigrette .60

DESSERTS

Gâteau moka .50	Cerises juteuses 1.25
Meringues glacées .60	Crêpes à la gelée .80
Crêpes Suzette 1.25	Crêpes nature .70
Glace sauce chocolat .60	Omelette au rhum 1.10
Fruits de saison à l'eau-de-vie .75	Glace à la vanille .50
Omelette soufflée à la Jules César (2) 2.00	Fraises au kirsch
Omelette Alaska Antoine (2) 2.50	Pêche Melba

FROMAGES

Roquefort .50	Liederkranz .50	Gruyère .50
Camembert .50	Fromage à la crème Philadelphie .50	

CAFÉ ET THÉ

Café .20	Café au lait .20	Thé .20
Café brûlé diabolique 1.00	Thé glacé .20	Demi-tasse

EAUX MINÉRALES—BIÈRE—CIGARES—CIGARETTES

White Rock	Cliquot Club	Bière locale	Canada Dry	Cigarettes
Vichy				

Roy B. Alciatore, Propriétaire

718-717 Rue St. Louis

Nouvelle Orléans, Louisiane

2

El Mítico Hombre-Mes

La buena cocina toma tiempo. Si usted está dispuesto a esperar, es para servirle mejor y complacerlo.

MENÚ DEL RESTAURANTE ANTOINE, NEW ORLEANS

Más proyectos de software han fracasado por la carencia de un calendario que debido a la combinación de otras causas. ¿Por qué la causa más común de este desastre resulta ser el calendario?

Primero, nuestras técnicas de estimación están poco desarrolladas. Y más grave aún es que reflejan una suposición que no se menciona y que es bastante falsa, i.e., que todo irá bien.

Segundo, nuestras técnicas de estimación confunden de forma equivocada el esfuerzo con el progreso, y ocultan esa suposición de que los hombres y los meses son intercambiables.

Tercero, debido a la incertidumbre de nuestras estimaciones, los gestores de software carecen de la amable tenacidad del cocinero del Antoine.

Cuarto, el calendario de avances está mal supervisado. En la ingeniería de software se consideran innovaciones radicales a técnicas comprobadas y habituales en otras disciplinas de la ingeniería.

Quinto, cuando se detecta una demora en el calendario, la respuesta normal (y típica) es añadir mano de obra. Esto es como apagar el fuego con gasolina, empeora mucho más las cosas. Más fuego exige más gasolina, y así se inicia un ciclo regenerativo con final desastroso.

La supervisión del calendario es materia de un ensayo aparte. Consideremos otros aspectos del problema con mayor detalle.

Optimismo

Todos los programadores son optimistas. Quizá sea porque este embrujo moderno atrae en especial a aquellos que creen en finales felices y hadas madrinas. O quizá tantas pequeñas decepciones ahuyentan a todos excepto a aquellos que habitualmente se concentran en el objetivo final. O a lo mejor es simplemente porque las computadoras son jóvenes, los programadores son más jóvenes, y los jóvenes son siempre optimistas. Pero sin embargo, el proceso de selección funciona, el resultado es irrefutable: “Esta vez con seguridad funcionará,” o “Acabo de encontrar el último error.”

Así pues, la primera falsa suposición que subyace al calendario de la programación de sistemas es que *todo irá bien*, i.e., que *cada tarea tardará solo el tiempo que “deba” hacerlo*.

La omnipresencia del optimismo entre los programadores merece más que un simple análisis. Dorothy Sayers, en su excelente libro, *La Mente del*

Creador, divide el trabajo creativo en tres etapas: la idea, la implementación, y la interacción. Así, un libro, o una computadora, o un programa llega a existir primero como una construcción ideal, elaborada fuera del tiempo y del espacio, aunque íntegra en la mente del autor. Se produce en el tiempo y el espacio, con pluma, tinta y papel, o a través de alambres, silicio y ferrita. La creación concluye cuando alguien lee el libro, usa la computadora, o ejecuta el programa, interactuando así con la mente del creador.

Esta descripción, que la Sra. Sayers usa para iluminar no solo el quehacer creativo sino también la doctrina cristiana de la trinidad, nos ayudará en la presente tarea. Para el ser humano hacedor de cosas, las incompletitudes e inconsistencias de nuestras ideas solo se aclaran durante la implementación. Así es como la escritura, la experimentación y “el desarrollo” son prácticas esenciales para el teórico.

En muchos trabajos creativos el medio de ejecución es inmanejable. Maderas partidas; manchas de pintura; cortos en circuitos eléctricos. Estas limitaciones físicas del medio restringen la expresión de las ideas, y además crean dificultades inesperadas a la hora de su implementación.

Así pues, la implementación toma tiempo y sudor debido tanto al medio físico como por lo inadecuado de nuestras ideas subyacentes. Tendemos a culpar al medio físico por la mayoría de nuestras dificultades de implementación; porque los medios no son “nuestros” en la forma en que nuestras ideas lo son, además nuestro orgullo tiende a empañar nuestro juicio.

Sin embargo, la programación de computadoras crea con un medio excesivamente maleable. El programador construye a partir del pensamiento puro: los conceptos y sus tan flexibles representaciones. Debido a que el medio es maleable, esperamos pocas dificultades en la implementación; de ahí nuestro optimismo omnipresente. Debido a que nuestras ideas son incorrectas, cometemos errores; por lo tanto, nuestro optimismo sea injustificado.

Dada una tarea, la suposición de que todo irá bien tiene un efecto probabilístico en el calendario. Y efectivamente la tarea podría ir como está planeada, pues encontraremos que existe una distribución de probabilidad para el retraso, y asimismo que “el no retraso” tiene una probabilidad finita. Sin embargo, un proyecto de programación grande consta de muchas tareas, algunas encadenadas de extremo a extremo. La probabilidad de que cada una vaya bien será cada vez más pequeña.

El Hombre-Mes

El segundo modo del pensamiento falaz se expresa en la propia unidad de esfuerzo utilizada en la estimación y calendarización: el hombre-mes. El costo en efecto varía como el producto del número de hombres y el número de meses. El progreso no. *Por lo tanto la unidad del hombre-mes como medida del tamaño de un trabajo es un mito peligroso y engañoso.* Pues implica que los hombres y los meses son intercambiables.

Los hombres y los meses son mercancías intercambiables solo cuando una tarea puede dividirse entre muchos trabajadores *sin comunicación entre ellos* (Fig. 2.1). Esto es cierto para la cosecha de trigo o la recolección de algodón; pero no es ni siquiera aproximadamente cierto en la programación de sistemas.



Fig. 2.1 Tiempo versus número de trabajadores – una tarea perfectamente divisible

Cuando una tarea no puede dividirse debido a las restricciones secuenciales, la aplicación de un mayor esfuerzo no tiene efecto sobre el calendario (Fig. 2.2). Dar a luz a un niño tarda nueve meses, sin importar cuantas mujeres se asignen. Muchas tareas de software tienen esta característica debido a la naturaleza secuencial de la depuración.

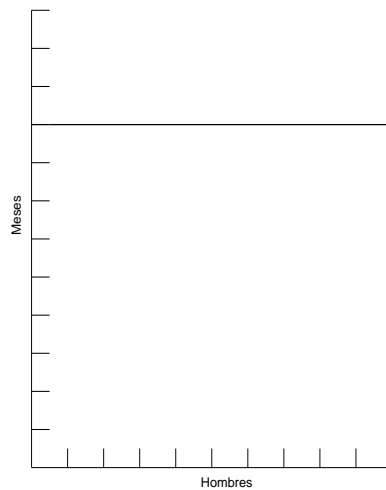


Fig. 2.2 Tiempo versus número de trabajadores – una tarea no divisible

En tareas que pueden dividirse pero que requieren comunicación entre subtareas, hay que añadir a la cantidad total de trabajo el esfuerzo de comunicación. Por lo tanto, lo mejor que podemos obtener es algo más pobre aún que el intercambio entre hombres y meses (Fig. 2.3).



Fig. 2.3 Tiempo versus número de trabajadores – una tarea divisible que requiere comunicación

La carga adicional de comunicación está compuesta de dos partes, entrenamiento e intercomunicación. Cada trabajador debe ser entrenado en la tecnología, los objetivos del proyecto, la estrategia global y el plan de trabajo. Este entrenamiento no se puede dividir, así que esta parte del esfuerzo añadido varía linealmente con el número de trabajadores.¹

La intercomunicación es peor. Si cada parte de la tarea debe coordinarse de forma separada con cada una de las otras partes el esfuerzo se incrementa como $n(n - 1)/2$. Tres trabajadores requieren tres veces más pares de intercomunicación que dos; cuatro trabajadores requieren seis veces más que dos. Si además, se necesitan reuniones de trabajo entre tres, cuatro, etc., trabajadores para resolver cosas juntos, el asunto todavía se pone peor. El esfuerzo adicional de comunicación puede contrarrestar totalmente la división de la tarea original y llevarnos a la situación de la Fig. 2.4.



Fig. 2.4 Tiempo versus número de trabajadores – una tarea con interrelaciones complejas

Puesto que la construcción del software es intrínsecamente un trabajo de sistemas – una operación de interrelaciones complejas – la labor de comunicación es grande, y domina rápidamente la reducción del tiempo de cada tarea, obtenida a través de la división. Por lo tanto, añadir más personas no reduce sino extiende el calendario.

Pruebas del Sistema

No hay parte del calendario más afectada por las restricciones secuenciales que la depuración de componentes y la prueba del sistema. Además, el tiempo requerido depende de la cantidad y lo escurridizo de los errores. Teóricamente este número debería ser cero. En virtud de nuestro optimismo, generalmente esperamos que el número de errores sea menor de lo que finalmente resulta ser. Por lo tanto, las pruebas generalmente representan la parte peor calendarizada de la programación.

Por algunos años he estado utilizando con éxito la siguiente regla empírica para calendarizar una tarea de software.

- 1/3 planificación
- 1/6 codificación
- 1/4 prueba de componentes y
- 1/4 pruebas del sistema con todos los componentes disponibles

Esta regla difiere de las calendarizaciones habituales en varios aspectos importantes:

1. La fracción dedicada a la planificación es mayor a la normal. Aun así, apenas es suficiente para realizar una especificación sólida y detallada, y no es suficiente para incluir la exploración o la investigación de técnicas totalmente nuevas.
2. La *mitad* del calendario está dedicada a la depuración de programas terminados y es mucho mayor a lo normal.
3. A la parte que es fácil de calcular, i.e., la codificación, se le asigna solo una sexta parte del calendario.

Al examinar proyectos calendarizados de forma tradicional, encontré que pocos asignaban una mitad del calendario a pruebas, aunque la mayoría en efecto invertía la mitad del calendario real para este propósito. Muchos de estos cumplían con el calendario hasta y exceptuando en las pruebas del sistema.²

En concreto, errar en asignar tiempo suficiente a pruebas del sistema es especialmente desastroso. Puesto que los retrasos vienen al final del calendario, nadie se da cuenta de los problemas hasta casi la fecha de entrega. Malas noticias, retrasado y sin aviso previo, eso es preocupante para clientes y gestores.

Además, el retraso en esta etapa tiene repercusiones financieras inusualmente severas, como también psicológicas. El personal del proyecto está totalmente completo, y el costo por día es máximo. Más grave aún, el software es para apoyar otros emprendimientos comerciales (el envío de computadoras, la operación de nuevas instalaciones, etc.) y los costos secundarios por el retraso son muy altos, pues ya casi es hora del envío del software. Y efectivamente, estos costos secundarios pueden sobrepasar por mucho a los demás. Por lo tanto, es muy importante disponer de tiempo suficiente para pruebas del sistema en el calendario original.

Estimación sin Agallas

Note que para el programador, como para el cocinero, la urgencia del cliente puede guiar la conclusión de la tarea, pero no puede guiar su conclusión real. Una tortilla, prometida en dos minutos, puede parecer una muy buena mejora. Pero cuando no ha cuajado en dos minutos, el cliente tiene dos opciones – esperar o comerlo crudo. Los clientes del software tienen las mismas opciones.

El cocinero tiene todavía otra opción más; puede aumentar el fuego. El resultado con frecuencia será una tortilla que nada podrá salvar – quemada en una parte, y cruda en otra.

Ahora bien, no creo que los gestores de software tengan menos coraje propio y firmeza que los cocineros, ni que otros gestores de la ingeniería. Pero los falsos calendarios con tal de cumplir con la fecha que desea el cliente es mucho más común en nuestra disciplina que en otro lugar de la ingeniería. Es muy difícil hacer una defensa vigorosa, convincente y con el riesgo de perder el trabajo de una estimación obtenida a través de un método no cuantitativo, respaldada por pocos datos, y apoyada principalmente por las ocurrencias de los gestores.

Claramente necesitamos dos soluciones. Tenemos que desarrollar y difundir estadísticas de productividad, estadísticas de incidencia de errores, reglas de estimación, etcétera. La profesión entera solo puede beneficiarse compartiendo tales datos.

Hasta que la estimación se asiente sobre bases sólidas, los gestores deben reforzar sus agallas y defender sus estimaciones con la seguridad de que sus modestas intuiciones son mejores que las estimaciones derivadas de sus deseos.

Desastre Regenerativo del Calendario

¿Qué hacer cuando un proyecto de software imprescindible está retrasado? Añadir fuerza laboral, naturalmente. Como sugieren las figuras de la 2.1 a la 2.4, esto puede ayudar o no.

Veamos un ejemplo.³ Supongamos que una tarea está estimada en 12 hombres-mes y se asignan tres personas por cuatro meses, y además existen hitos medibles A, B, C, D, que están calendarizados para caer al final de cada mes (Fig. 2.5).

Ahora suponga que no se cumplió la primera etapa hasta que transcurrieron dos meses (Fig. 2.6). ¿Cuáles son las alternativas que encara el gestor?

1. Suponga que la tarea debe terminarse a tiempo. Suponer que solo la primera parte de la tarea fue mal estimada, es la historia que cuenta exactamente la Fig. 2.6. Entonces restan 9 hombres-mes de esfuerzo, y dos meses, así que se necesitarán $4\frac{1}{2}$ personas. Hay que añadir 2 personas a las 3 ya asignadas.
2. Suponga que la tarea debe terminarse a tiempo. Suponer que toda la estimación fue mala en general, la Fig. 2.7 describe realmente la situación. Por lo tanto, restan 18 hombres-mes de esfuerzo, y dos meses, así que se necesitarán 9 personas. Hay que añadir 6 personas a las 3 ya asignadas.

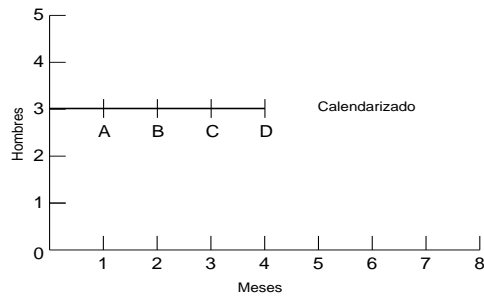


Figura 2.5

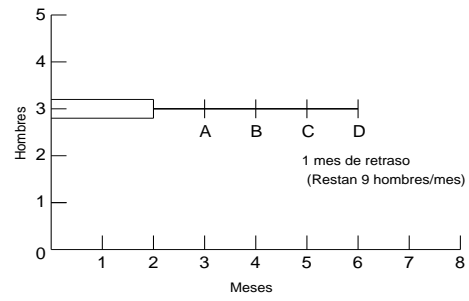


Figura 2.6

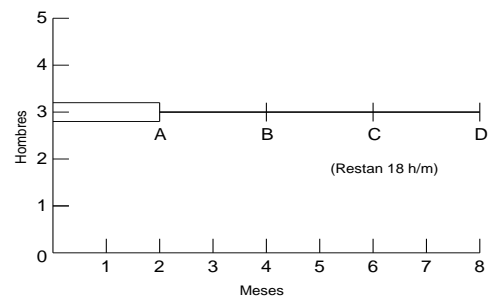


Figura 2.7

3. Recalendarice. Me gusta el consejo dado por P. Fagg, un experimentado ingeniero de hardware, “No cometa errores pequeños.” Eso significa, permitir tiempo suficiente en el nuevo calendario para asegurar que el trabajo se lleve a cabo completa y cuidadosamente, y no se tenga que recalendarizar de nuevo.
4. Recorte la tarea. En la práctica esto tiende a suceder de cualquier manera, una vez que el equipo observa la demora en el calendario. Esta es la única acción viable cuando los costos secundarios por el retraso son muy elevados. Las únicas alternativas del gestor son: recortar la tarea formal y cuidadosamente, para recalendarizar, u observar que la tarea se recorte silenciosamente mediante un diseño precipitado y pruebas incompletas.

En los primeros dos casos, insistir en que la tarea inalterada sea concluida en cuatro meses es desastroso. Por ejemplo, considere los efectos regenerativos en la primera alternativa (Fig. 2.8). Las dos personas nuevas, que a pesar de ser competentes y haber sido reclutadas a prisa, requerirán ser entrenadas en la tarea por una persona con experiencia. Si esto toma un mes, *3 hombres-mes habrán sido dedicados a trabajar pero no en la estimación original*. Además, la tarea dividida originalmente en tres partes, debe dividirse nuevamente en cinco partes: por lo tanto, se perderá algo del trabajo ya realizado, y las pruebas del sistema deberán ser prolongadas. Así, al final del tercer mes, restan en esencia más de 7 hombres-mes de trabajo, y se dispone de 5 personas entrenadas y un mes. Como la Fig. 2.8 sugiere, el producto está retrasado como si nadie hubiese sido añadido (Fig. 2.6).

El deseo de llevarlo a cabo en cuatro meses, considerando solo el tiempo de entrenamiento sin ninguna repartición ni pruebas extras de los sistemas, requeriría añadir al final del segundo mes, 4 personas, no 2. Para cubrir los efectos de la repartición y las pruebas del sistema, se tendría que añadir otra persona más. Sin embargo, ahora al menos se tiene un equipo de 7 personas, no uno de 3; así pues, tales aspectos como la organización y la división de tareas son diferentes en tipo, no solo en grado.

Note que por el final de tercer mes el panorama es sombrío. El hito del 1

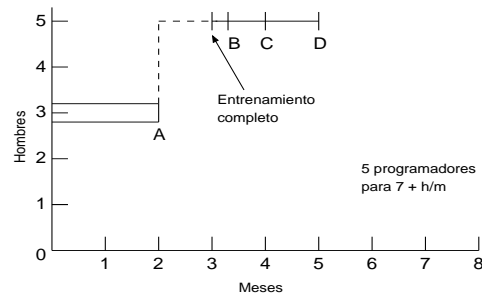


Figura 2.8

de Marzo no ha sido alcanzado a pesar de todo el trabajo de gestión. La tentación de repetir el ciclo es muy fuerte, añadir aún más mano de obra. Ahí yace la locura.

Lo anterior suponía que solo el primer hito fue mal estimado. Si en Marzo se hace la suposición conservadora de que todo el calendario fue optimista, como muestra la Fig. 2.7, se busca añadir 6 personas solo a la tarea original. El cálculo de los efectos del entrenamiento, de la repartición y las pruebas del sistema se dejan al lector como ejercicio. Sin duda alguna, el desastre regenerativo conducirá a un producto más pobre, más retrasado, que si se hubiera recalendarizado la tarea con las tres personas originales, sin añadir a nadie.

Sobresimplificando exageradamente, enunciamos la Ley de Brooks:

Añadir mano de obra a un proyecto de software retrasado lo retrasa aún más.

Pues esta es la desmitificación del hombre-mes. El número de meses de un proyecto está sujeto a restricciones secuenciales. El máximo número de personas está sujeto al número de subtareas independientes. A partir de estas dos cantidades podemos obtener calendarios que usen menos personas y más meses. (El único riesgo es la obsolescencia del producto.) Sin embargo, no

se puede obtener un calendario capaz de funcionar usando más personas y menos meses. Gran parte de los proyectos de software han fracasado más por la carencia de un calendario que debido a la combinación de otras causas.

3

El Equipo Quirúrgico



3

El Equipo Quirúrgico

Estos estudios han revelado que las diferencias entre individuos de alto y bajo rendimiento, generalmente son de un orden de magnitud.

SACKMAN, ERIKSON, Y GRANT¹

Foto UPI/El Archivo Bettman

En reuniones de la sociedad de computación, con frecuencia escucho a jóvenes gestores de programación afirmar que están a favor de un equipo pequeño e inteligente; de personal de primera clase, en lugar de un proyecto con cientos de programadores, y por ende mediocres. Nosotros también estamos a favor.

Aunque esta ingenua exposición de alternativas evade el meollo del problema – ¿cómo construir *sistemas grandes* en base a calendarios sensatos? Veamos cada lado de esta pregunta con mayor detalle.

El Problema

Hace mucho que los gestores de programación han reconocido la amplia variación entre los buenos y los malos programadores. Aunque nos ha asombrado a todos la magnitud de la medición real. En uno de sus estudios, Sackman, Erikson, y Grant, midieron el rendimiento de un grupo de experimentados programadores. Solo dentro de este grupo, la razón promedio del rendimiento entre los mejores y los peores programadores era cerca de 10:1 en mediciones de productividad y un impresionante 5:1 en mediciones de velocidad y tamaño del programa! En resumen, el programador de 20,000 dólares/año bien podría ser 10 veces más productivo que uno de 10,000 dólares/año. Lo contrario también podría ser cierto. Los datos no mostraron ninguna correlación en absoluto entre experiencia y rendimiento. (Dudo que eso sea universalmente válido.)

Anteriormente he afirmado que la coordinación total de las mentes afecta el costo del trabajo, ya que una gran parte de éste es el costo de la comunicación y la corrección de los efectos adversos de la mala comunicación (depuración del sistema). Esto, también, sugiere que se busque que el sistema sea construido por el menor número de mentes posible. Y en efecto, gran parte de la experiencia con sistemas de programación grandes muestra que el enfoque de la fuerza bruta es costoso, lento, ineficiente, y produce sistemas que no están conceptualmente integrados. OS/360, Exec 8, Scope 6600, Multics, TSS, SAGE, etc. – y la lista continúa.

La conclusión es simple: si un proyecto de 200 personas tiene 25 gestores que son los programadores más competentes y experimentados, despida a la tropa de 175 y ponga a los gestores de nuevo a programar.

Ahora examinemos esta solución. Por un lado, fracasa en el enfoque ideal de un equipo inteligente y *pequeño*, el cual por consenso no debe exceder las

10 personas. Este equipo como es muy grande necesitará al menos dos niveles de gestión, o cerca de cinco gestores. Y adicionalmente necesitará apoyo financiero, personal, espacio, secretarías, y operadores de computadoras.

Por otro lado, el equipo original de 200 personas no era lo suficientemente grande como para construir sistemas realmente grandes a través de los métodos de la fuerza bruta. Por ejemplo, considere el OS/360. En el punto más alto trabajaban más de 1000 personas en él – programadores, escritores, operadores, empleados, secretarías, gestores, grupos de apoyo, y demás. Probablemente de 1963 a 1966 se dedicaron 5000 hombres-año a su diseño, construcción, y documentación. A nuestro equipo candidato de 200 personas, le tomaría 25 años llevar el producto a su etapa actual, si los hombres y los meses fueran todavía intercambiables!

Pues este es el problema con el concepto de equipo pequeño e inteligente: *es muy pequeño para sistemas realmente grandes*. Considere el trabajo del OS/360 como si pudiera ser abordado por un equipo pequeño e inteligente. Escoja un equipo de 10-personas. Como límite, ya que son inteligentes, permita que sean siete veces más productivos que los programadores mediocres, respecto a la programación y la documentación. Suponga que el OS/360 fue construido por programadores mediocres (lo cual está lejos de ser cierto). Suponga además, como límite otro factor de siete en el aumento de la productividad que viene de la reducción en la comunicación por parte del equipo pequeño. Asuma que el *mismo* equipo permanece a lo largo de todo el proyecto. Ahora bien, $5000/(10 \times 7 \times 7) = 10$, ellos podrán realizar el trabajo de 5000 hombres-año en 10 años. ¿Será un producto interesante 10 años después de su diseño inicial? ¿O se habrá hecho obsoleto debido al acelerado desarrollo de la tecnología del software?

El dilema es cruel. Para la eficiencia e integridad conceptual, es preferible que unas pocas mentes realicen el diseño y la construcción. Sin embargo, para los sistemas grandes se busca añadir considerable mano de obra de apoyo, tal que el producto se pueda presentar a tiempo. ¿Cómo poder reconciliar estos dos requisitos?

La Propuesta de Mills

La propuesta de Harlan Mills ofrece una solución nueva y creativa.^{2,3} Él propone que cada segmento de un trabajo grande sea abordado por un equipo,

y que cada equipo sea organizado como un equipo quirúrgico en vez de un equipo de carnicería. Esto es, en lugar de que cada miembro haga cortes, solo uno hace los cortes y los demás le brindan apoyo para que mejore su eficacia y su productividad.

Si reflexionamos un poco vemos que este concepto cumple con la desiderata, si podemos lograr que funcione. Pocas mentes se involucran en el diseño y la construcción, aunque se usan muchas manos para ponerlo en marcha. ¿Funcionará? ¿Quién será el anestesiólogo? y ¿Quiénes serán las enfermeras en el equipo de programación? y ¿Cómo se dividirá el trabajo? Permítanme mezclar metáforas libremente y sugerir cómo tal equipo podría funcionar si lo extendemos para incluir todo el apoyo posible.

El cirujano. Mills lo llama *programador líder*. Él personalmente define las especificaciones tanto funcionales como de rendimiento, diseña los programas, los codifica, los prueba, y escribe su documentación. Escribe en un lenguaje de programación estructurado como PL/I, y tiene un acceso eficaz al sistema de cómputo que no solo corre sus pruebas sino también almacena varias versiones de sus programas, permitiendo una fácil actualización de archivos, y le provee un editor de texto para que lleve a cabo la documentación. El cirujano debe tener un gran talento, diez años de experiencia, y un considerable conocimiento de sistemas y aplicaciones, ya sea en matemáticas aplicadas, en manejo de datos de negocios, o algo semejante.

El copiloto. Es el alter ego del cirujano, es capaz de realizar cualquier parte del trabajo, aunque es menos experimentado. Su función principal es participar en el diseño como un pensador, un debatidor y un evaluador. El cirujano le sugiere ideas, aunque no está sujeto a sus consejos. Normalmente, el copiloto es el representante de su equipo en discusiones con otros equipos acerca del funcionamiento y las interfaces. Conoce todo el código íntimamente. Investiga estrategias alternativas de diseño. Obviamente sirve como seguro del cirujano contra desastres. Puede incluso escribir código, aunque no es el responsable por ninguna parte del mismo.

El administrador. El cirujano es el jefe, y debe tener la última palabra acerca del personal, los aumentos, el espacio, y demás, aunque casi no debe invertir su tiempo en estas cuestiones. Por tanto, necesita un administrador profesional que maneje el dinero, el personal, el espacio, y las máquinas, y que sea la interfaz con la maquinaria administrativa del resto de la organi-

zación. Baker sugiere que el administrador tenga un trabajo de tiempo completo solo si el proyecto tiene considerables requisitos legales, contractuales, de informes, o financieros debido a la relación usuario-productor. De otra manera, un solo administrador puede servir a dos equipos.

El editor. El cirujano es el responsable de generar la documentación – para una máxima claridad es necesario que él la escriba. Esto se aplica tanto para las descripciones internas y como para las externas. Sin embargo, el editor toma el borrador o el manuscrito dictado por el cirujano y lo analiza, lo reelabora, le añade referencias y bibliografía, lo mantiene a través de sus distintas versiones, y supervisa el proceso de producción.

Dos secretarías. El administrador y el editor necesitarán cada uno una secretaria: la secretaria administrativa manejará la correspondencia del proyecto y los archivos que no sean técnicos.

El archivista de programas. Es el responsable del mantenimiento de todos los registros técnicos del equipo dentro de una biblioteca de productos de programación. El archivista está entrenado como una secretaria y es el responsable de los archivos legibles tanto por la computadora como por los humanos.

Todo ingreso a la computadora va al archivista, quien lo registra y tipea en caso necesario. Asimismo, archiva e indexa los listados de salida. Las ejecuciones más recientes de cualquier modelo se mantienen en una libreta de estado; mientras que todas las anteriores se almacenan en un archivo cronológico.

Algo absolutamente vital al concepto de Mills es la transformación de la programación “de un arte privado a una práctica pública” pues logra hacer claras *todas* las operaciones de la computadora a todos los miembros del equipo y señala que todos los programas y datos son propiedad del equipo, y no propiedad privada.

La labor especializada del archivista releva al programador de las tareas de oficina, sistematiza y asegura un desempeño adecuado en esas tareas que con frecuencia se eluden, y mejora el activo más valioso del equipo – el producto del trabajo. Claramente el concepto expuesto anteriormente asume ejecuciones por lotes. Cuando se usan terminales interactivas, particularmente aquellas con salidas no impresas, las funciones del archivista no disminuyen, sino que cambian. Ahora registra todas las actualizaciones de las

copias de los programas del equipo a partir de copias de trabajo privadas, todavía se encarga de las ejecuciones por lotes, y usa sus propios medios interactivos para controlar la integridad y disponibilidad del producto en desarrollo.

El especialista en herramientas. Hoy es muy fácil contar con servicios de edición de archivos, edición de texto, y depuración interactiva, de tal manera que un equipo rara vez necesitará su propia máquina y personal de operadores. Estos servicios deben estar disponibles con una respuesta y una fiabilidad indiscutiblemente satisfactorias; y el cirujano debe ser el único juez de la idoneidad del servicio puesto a su disposición. Aparte, necesitará un especialista en herramientas, responsable de asegurar la adecuación del servicio básico y de la construcción, mantenimiento, y actualización de herramientas especiales que el equipo necesita – en su mayoría serán servicios de cómputo interactivo. Cada equipo necesitará su propio especialista en herramientas, independientemente de la calidad y confiabilidad de cualquier servicio proporcionado por la administración central, porque su trabajo es velar por las herramientas que su cirujano necesita o busca, sin tener en cuenta las necesidades de los demás equipos. El fabricante de herramientas a menudo construirá utilerías especializadas, catálogo de procedimientos, y bibliotecas de macros.

El probador. El cirujano necesitará un adecuado banco de casos de prueba para probar piezas de su trabajo mientras los escribe, y después para la prueba general. Por lo tanto, el probador es por un lado un adversario que idea casos de prueba del sistema a partir de las especificaciones funcionales, y por otro lado, un colaborador que idea datos de prueba para la depuración diaria. También podría planificar secuencias de pruebas y construir el andamiaje necesario para pruebas de componentes.

El abogado del lenguaje. Cuando apareció Algol, el personal empezó a reconocer que la mayoría de las instalaciones de computadoras tenían una o dos personas que se deleitaban con el dominio de las complejidades de los lenguajes de programación. Estos expertos resultaron ser muy útiles y ampliamente consultados. Este talento es bastante diferente del talento del cirujano, quien es principalmente un diseñador de sistemas y alguien que idea representaciones. El abogado del lenguaje puede encontrar una manera pulcra y eficiente de usar el lenguaje para realizar cosas difíciles, oscuras, o truculentas. A menudo deberá hacer estudios breves (de dos o tres días) acerca de una

buena técnica. Un abogado del lenguaje puede servir a dos o tres cirujanos.

Pues así es como 10 personas pueden contribuir en roles bien diferenciados y especializados en un equipo de programación construido bajo el modelo quirúrgico.

Cómo funciona

El equipo organizado de esta manera cumple con la desiderata de varias formas. Diez personas, siete de ellas profesionales, trabajan en un problema, pero el sistema es el producto de una sola mente – o a lo mucho dos, actuando *uno animo*.

En particular, observe las diferencias entre un equipo de dos programadores organizados convencionalmente y el equipo cirujano-copiloto. Primero, en el equipo convencional los participantes se dividen el trabajo, y cada uno es responsable del diseño e implementación de una parte del trabajo. En el equipo quirúrgico, el cirujano y el copiloto están cada uno al tanto de todo el diseño y todo el código. Esto ahorra el trabajo de la asignación de espacio de memoria, de los accesos a disco etc. Y también asegura la integridad conceptual del trabajo.

Segundo, en el equipo convencional los participantes son iguales, las inevitables diferencias deben ser discutidas o negociadas. Puesto que el trabajo y los recursos están divididos, las diferencias de opinión están confinadas a la estrategia global y a la interfaz, aunque estén compuestas por diferencias de interés – e.g., a quienes cuyo espacio de memoria será usado para un búfer. En el equipo quirúrgico, no hay diferencias de interés, y las diferencias de opinión las resuelve el cirujano unilateralmente. Estas dos diferencias – la carencia de división del problema y la relación superior-subordinado hacen posible que el equipo quirúrgico actúe *uno animo*.

Incluso la especialización de funciones del resto del equipo es la clave para su eficiencia, y permite un patrón de comunicación radicalmente más simple entre sus miembros, como muestra la Fig. 3.1.



Fig. 3.1 Patrón de comunicación en equipos de programación de 10 personas

El artículo de Backer³ reporta acerca de una sola prueba a pequeña escala del concepto de equipo. Para ese caso funcionó como se predijo, con resultados estupendamente buenos.

A Mayor Escala

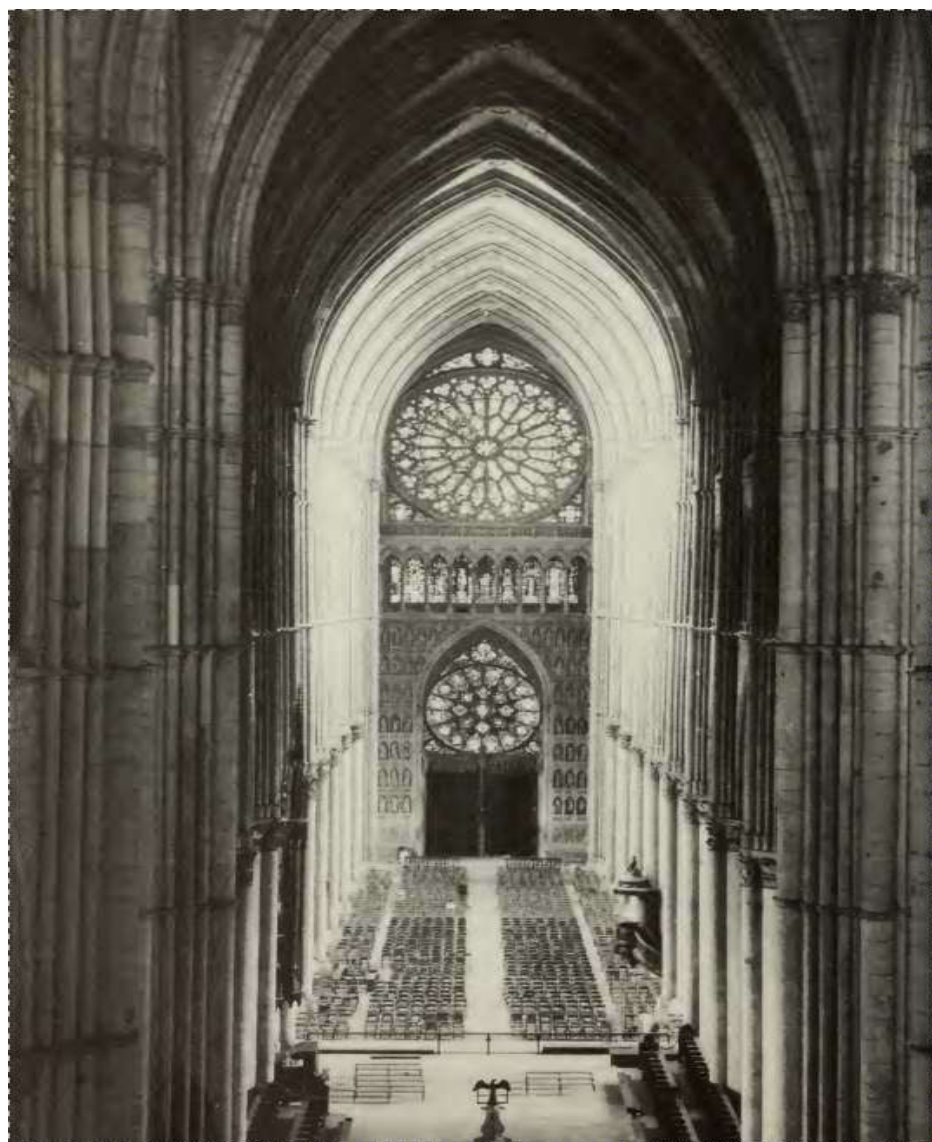
Al menos por el momento. Sin embargo, el problema es cómo construir cosas que tomen 5000 hombres-año, no cosas que tomen 20 o 30. Un equipo de 10 personas puede ser eficaz sin importar cómo se organice, siempre que *todo* el trabajo esté dentro de su ámbito. Pero ¿Cómo se usa el concepto de equipo quirúrgico en trabajos grandes cuando se involucran varios cientos de personas para llevar a cabo la tarea?

El éxito del proceso a mayor escala depende del hecho de que la integridad conceptual de cada pieza haya sido mejorada radicalmente – ese número de mentes que determinan el diseño se ha dividido entre siete. Así es posible poner 200 personas en un problema y encarar el problema de coordinar solo 20 mentes, las de los cirujanos.

Sin embargo, para el problema de coordinación se deben usar técnicas distintas, que son discutidas en los siguientes capítulos. Por el momento es suficiente decir que todo el sistema también debe tener integridad conceptual, y que se necesita un arquitecto del sistema que lo diseñe completamente, de forma descendente. Para hacer el trabajo manejable, es importante distinguir claramente entre arquitectura e implementación, el arquitecto del sistema debe restringirse él mismo escrupulosamente a la arquitectura. De cualquier manera, tales roles y técnicas han mostrado ser viables y, asimismo, muy productivas.

4

*Aristocracia, Democracia, y
Diseño de Sistemas*



4

Aristocracia, Democracia, y Diseño de Sistemas

Esta gran iglesia es una obra de arte incomparable. No existe aridez, ni confusión en los principios expuestos...

Es el cenit de un estilo, el trabajo de artistas que entendieron y asimilaron todos los éxitos de sus predecesores, aunque en completa posesión de las técnicas de su tiempo, las usaron sin la indiscreta exposición ni la injustificada proeza del talento.

Este fue Jean d'Orbais quien sin duda concibió el plan general del edificio, un plan que ha sido respetado, al menos en sus elementos esenciales, por sus sucesores. Esta es una de las razones por la extrema coherencia y unidad del edificio

GUÍA DE LA CATEDRAL DE REIMS¹

Fotografías de Emmanuel Boudot-Lamotte

Integridad Conceptual

La construcción de la mayoría de las catedrales Europeas muestran diferencias en sus diversas partes en el plan o en el estilo arquitectónico, pues fueron realizadas por diferentes constructores y en distintas épocas. Los últimos constructores fueron tentados para “mejorar” el diseño de los anteriores, para reflejar por un lado los cambios en la moda y por otro las diferencias en el gusto personal. Así, la tranquilidad del crucero Normando colinda y contradice la elevada nave Gótica, y el resultado proclama el orgullo de los constructores tanto como la gloria de dios.

Contra estos, la unidad arquitectónica de Reims se yergue en un espléndido contraste. El placer que emociona a los espectadores proviene tanto de la integridad del diseño como de cualquiera de sus méritos particulares. Como menciona la guía, la integridad se llevó a cabo a través de la autoabnegación de ocho generaciones de constructores, cada una de las cuales sacrificó algo de sus propias ideas con tal de que el todo pudiera ser un diseño puro. El resultado proclama no sólo la gloria de dios, sino también su poder para salvar a los pecadores de su soberbia.

Aun cuando no dispongamos de siglos para construir, la mayoría de los sistemas de programación reflejan una desunión conceptual mucho peor que las catedrales. Por lo general, esto no surge a partir de una sucesión en serie de diseñadores expertos, sino de la separación del diseño en muchas tareas llevadas a cabo por muchas personas.

Yo sostengo que la integridad conceptual es *la* consideración más importante en el diseño de sistemas. Es mejor tener un sistema que omita ciertos rasgos y mejoras anómalos, pero que refleje un conjunto de ideas de diseño, que tener uno que contenga muchas ideas buenas pero independientes y sin coordinación. En este capítulo y en los dos siguientes, examinaremos las consecuencias de este tema para el diseño de los sistemas de programación:

- ¿Cómo se logra la integridad conceptual?
- ¿Este argumento acaso no implica tener una élite, o una aristocracia de arquitectos, y una horda de implementadores plebeyos cuyos talentos creativos e ideas han sido suprimidos?
- ¿Cómo se evita que los arquitectos vuelen alto con especificaciones no

implementables o costosas?

- ¿Cómo se garantiza que cada detalle insignificante de una especificación arquitectónica sea comunicado al implementador, sea entendido correctamente por él, e incorporado exactamente en el producto?

Lograr la Integridad Conceptual

El propósito de los sistemas de programación es construir computadoras que sean fáciles de usar. Para lograr esto, se suministran lenguajes y diversos medios que de hecho son programas invocados y controlados por las propias características del lenguaje. Pero estos programas se compran a cierto precio: la descripción externa de un sistema de programación es de diez a veinte veces mayor que la descripción externa del propio sistema de cómputo. Así, el usuario encontrará que es mucho más fácil especificar cualquier función en particular, aunque hay mucho más de donde escoger, y muchas más opciones y formatos que recordar.

La facilidad de uso se mejora solo si el tiempo ganado en la especificación funcional excede el tiempo perdido en aprender, recordar, y buscar manuales. Con los sistemas de programación modernos esta ganancia supera el costo, aunque en los años recientes la proporción de ganancia a costo parece haber disminuido conforme se fueron agregando funciones más complejas. Estoy impresionado por la facilidad de uso de la memoria de la IBM 650, incluso sin un programa ensamblador o ningún otro tipo de software.

Como el propósito es la facilidad de uso, la proporción entre funcionalidad y complejidad conceptual es la prueba final del diseño de sistemas. Ni la funcionalidad, ni la simplicidad por sí solas definen un buen diseño.

Este punto se ha malentendido ampliamente. El Operative System/360 es aclamado por sus constructores como la máquina más fina jamás construida, debido a que indiscutiblemente tiene la mayor funcionalidad. Pues la medida de excelencia de sus diseñadores fue siempre la funcionalidad, no la simplicidad. Por otro lado, el Sistema de Tiempo-Compartido para la PDP-10 es aclamado por sus constructores como el más fino, debido a su simplicidad y la sobriedad de sus conceptos. Sin embargo, por cualquier medición su funcionamiento no está ni siquiera en la misma clase que el OS/360. En cuanto

tomamos la facilidad de uso como criterio, todos ellos parecen estar desbalanceados, y solo alcanzan a cumplir la mitad de sus verdaderas metas.

Sin embargo, para un cierto nivel de funcionalidad, el mejor sistema es aquel en el cual se pueden especificar cosas con la mayor simplicidad y sencillez posibles. Pero la *simplicidad* no es suficiente. El lenguaje TRAC de Mooers y Algol 68 logran la simplicidad medida por el número de conceptos elementales distintos. Sin embargo, no son *sencillos*. Pues, para expresar cosas que con frecuencia uno quiere hacer se requieren combinaciones complicadas e imprevistas de los instrumentos básicos. No es suficiente aprender los elementos y las reglas para combinarlas; también se debe aprender el uso idiomático, y toda una tradición de cómo se combinan los elementos en la práctica. La simplicidad y la sencillez provienen de la integridad conceptual. Cada parte debe reflejar la misma filosofía y el mismo equilibrio de desiderata. Incluso cada parte debe usar las mismas técnicas en la sintaxis y nociones análogas en la semántica. Por lo tanto, la facilidad de uso impone la unidad del diseño y la integridad conceptual.

Aristocracia y Democracia

La integridad conceptual a su vez obliga a que el diseño proceda de una sola mente, o de un muy pequeño número de mentes en sintonía.

Sin embargo, las presiones del calendario fuerzan a que la construcción del sistema requiera muchas manos. Se dispone de dos técnicas para resolver este dilema. La primera es una cuidadosa división del trabajo entre arquitectura e implementación. La segunda es la nueva forma de estructurar los equipos encargados de la implementación de la programación discutida en el capítulo anterior.

Una forma muy poderosa de lograr la integridad conceptual en proyectos muy grandes consiste en separar la arquitectura de la implementación. Yo mismo lo he visto aplicado con gran éxito en la computadora Stretch de IBM y en la línea de productos de la computadora System/360. Y también su fracaso por no aplicarlo en la Operative System/360.

Por *arquitectura* de un sistema, quiero decir la especificación completa y detallada de la interfaz de usuario. Para una computadora es el manual de programación. Para un compilador es el manual del lenguaje. Para un programa de control son los manuales del lenguaje o los lenguajes utilizados para

invocar sus funciones. Para el sistema completo es la unión de los manuales que el usuario debe consultar para realizar toda su labor.

El arquitecto de un sistema, como el arquitecto de un edificio, es el representante del usuario. Su trabajo es brindar un conocimiento profesional y técnico para apoyar el genuino interés del usuario, opuesto al interés del vendedor, fabricante, etc.²

Debemos distinguir cuidadosamente la arquitectura de la implementación. Como Blaauw afirma, “Donde la arquitectura dice *qué* sucede, la implementación dice *cómo* se hace para que suceda.”³ Da como un ejemplo sencillo el reloj, cuya arquitectura consta de la carátula, las manecillas, y el botón de la cuerda. Cuando un niño ha aprendido esta arquitectura, puede decir la hora tan fácilmente desde un reloj de pulsera hasta una torre de iglesia. Sin embargo, la implementación y su realización, describen qué ocurre dentro del empaque – impulsado y controlado con exactitud por una variedad de mecanismos.

Por ejemplo, en la System/360 una sola arquitectura de computadora se implementó de forma muy diferente en cada uno de los nueve modelos. Al contrario, una sola implementación del flujo de datos, la memoria, y el microcódigo de la Model 30, se usaron en distintos momentos para cuatro arquitecturas diferentes: una computadora System/360, un canal múltiple con hasta 224 subcanales lógicamente independientes, un canal selector, y una computadora 1401.⁴

La misma distinción es igualmente aplicable a los sistemas de programación. Existe un Fortran IV estándar. Esta es la arquitectura para muchos compiladores. Muchas implementaciones son posibles dentro de esta arquitectura: texto en el núcleo o compilador en el núcleo, compilación rápida u optimización, dirigida por sintaxis o *ad-hoc*. Del mismo modo, cualquier lenguaje ensamblador o lenguaje de control de tareas admite muchas implementaciones del ensamblador o del planificador.

Ahora podemos abordar ese asunto tan sensible de la aristocracia contra la democracia. ¿No son los arquitectos una nueva aristocracia, una élite intelectual, creada para decirles a esos pobres implementadores tontos qué hacer? ¿No ha sido secuestrado todo el trabajo creativo por esta élite, dejando a los implementadores como engranajes de una máquina? ¿No obtendríamos un mejor producto recogiendo las buenas ideas de todo el equipo, siguiendo una filosofía democrática, en lugar de restringir el desarrollo de las especifi-

caciones a unos cuantos?

Respecto a la última pregunta, es la más fácil. Ciertamente no voy a sostener que sólo los arquitectos tienen buenas ideas arquitectónicas. Con frecuencia, los conceptos nuevos vienen de un implementador o de un usuario. Sin embargo, toda mi propia experiencia me convence, y he tratado de demostrar, que la integridad conceptual de un sistema determina su facilidad de uso. Las buenas ideas y características que no se integran con los conceptos básicos del sistema es mejor omitirlas. Si aparecen muchas de tales ideas importantes pero incompatibles, se desecha todo el sistema y se empieza de nuevo en un sistema integrado con conceptos básicos diferentes.

Como cargo a la aristocracia, la respuesta debe ser sí y no. Sí, en el sentido de que deben ser pocos arquitectos, su producto debe perdurar más que el de un implementador, y el arquitecto se coloca en el foco de las fuerzas que, en última instancia, se deben resolver en el interés del usuario. Si un sistema es tener integridad conceptual, alguien debe controlar los conceptos. Esa es una aristocracia que no necesita justificación.

No, porque el establecer especificaciones externas no es un trabajo más creativo que el diseño de las implementaciones. Es solo un trabajo creativo diferente. El diseño de una implementación, dada una arquitectura, requiere y permite tanto diseño creativo, como muchas ideas nuevas, y tanto brillantez técnica como diseño de las especificaciones externas. Asimismo, la relación costo-rendimiento del producto dependerá en gran medida del implementador, así como la facilidad de uso depende en gran medida del arquitecto.

Existen muchos ejemplos de otras artes y oficios que nos llevan a creer que la disciplina es buena para el arte. En efecto, el aforismo de un artista dice, "La forma es liberadora." Los peores edificios son aquellos cuyos presupuestos eran demasiado grandes para los propósitos que servían. La creatividad de Bach apenas parece haber sido acallada por la necesidad de producir una cantata de forma limitada cada semana. Estoy seguro de que la computadora Stretch hubiera tenido una mejor arquitectura si hubiera estado más restringida. Las restricciones impuestas por el presupuesto a la System/360 Model 30 fueron, en mi opinión, totalmente beneficiosas para la arquitectura de la Model 75.

Igualmente, he visto que el suministro externo de una arquitectura mejora, no limita, el estilo creativo del grupo de implementación. Es más, éste de

inmediato se concentra en la parte del problema que nadie ha abordado, y los inventos empiezan a fluir. Dentro de un grupo de implementadores sin restricciones, la mayoría de las ideas y debates se centran en las decisiones arquitectónicas, y la implementación en sí recibe poca atención.⁵

Este efecto, que lo he visto muchas veces, fue confirmado por R. W. Conway, cuyo grupo en Cornell construyó el compilador PL/C para el lenguaje PL/I. Él afirma, “Finalmente decidimos implementar el lenguaje sin cambios ni mejoras, porque los debates acerca del lenguaje hubieran agotado todo nuestro esfuerzo.”⁶

¿Mientras Tanto Qué Hacen los Implementadores?

Cometer un error multimillonario es una experiencia muy humillante, aunque también es muy memorable. Recuerdo claramente la noche que decidimos cómo organizar la escritura de las especificaciones externas para el OS/360. El director de la arquitectura, el director de la implementación del programa de control, y yo estuvimos machacando el plan, el calendario, y la división de responsabilidades.

El director de arquitectos tenía 10 hombres capaces. Afirmó que ellos podían escribir las especificaciones y hacerlo correctamente. Esto tomaría diez meses, tres más de lo que permitía el calendario.

El director del programa de control tenía 150 hombres. Afirmó que ellos podían preparar las especificaciones, en coordinación con el equipo de arquitectos; estaría bien hecho y sería práctico, y podía terminarlo a tiempo. Además, si el equipo de arquitectos lo hacía, sus 150 hombres estarían sentados meneando los dedos por diez meses.

A esto, el director de arquitectos respondió que si yo le daba la responsabilidad al equipo del programa de control, el resultado de hecho *no* estaría a tiempo, sino que estaría también retrasado tres meses, y de mucha peor calidad. Lo hice, y así fue. Estaba en lo cierto en ambas cuentas. Más aún, la falta de integridad conceptual hizo que el sistema fuera mucho más costoso de construir y cambiar, y diría que añadió un año al tiempo asignado a la depuración.

Por supuesto que se involucraron muchos factores en la decisión errada; pero el más abrumador fue el calendario y lo atractivo de poner a esos 150 im-

plementadores a trabajar. Es este canto de sirenas de cuyo riesgo fatal ahora soy consciente.

De hecho, cuando se propuso que un equipo pequeño de arquitectos escribiera todas las especificaciones externas para una computadora o sistema de programación, los implementadores plantearon tres objeciones:

- Las especificaciones serían bastantes ricas en funciones y no reflejarían las consideraciones prácticas del costo.
- Los arquitectos obtendrían toda la diversión creativa y excluirían la inventiva de los implementadores.
- La mayoría de los implementadores tendrían que sentarse ociosamente a que llegaran las especificaciones a través del estrecho embudo que era el equipo de arquitectos.

La primera de ellas es un peligro real, y será tratada en el próximo capítulo. Las otras dos son simples y puras fantasías. Como vimos anteriormente, la implementación es también una labor creativa de primer orden. La oportunidad de ser creativos e inventivos en la implementación no disminuye significativamente al trabajar dentro de unas especificaciones externas dadas, e incluso esa práctica puede mejorar la estructura de la creatividad. El producto completo sin duda lo hará.

La última objeción es de sincronización y coordinación. Una respuesta rápida es abstenerse de contratar implementadores hasta que las especificaciones estén completas. Esto se hace cuando se construye un edificio.

Sin embargo, en el negocio de los sistemas de computación el ritmo es más ágil, y uno desea comprimir el calendario lo más posible. ¿Cuánto se pueden traslapar las especificaciones y la construcción?

Como Blaauw señala, el esfuerzo creativo íntegro involucra tres fases distintas: arquitectura, implementación, y realización. De hecho, resulta que pueden empezar en paralelo y avanzar simultáneamente.

Por ejemplo, en el diseño de la computadora el implementador puede empezar tan pronto como tenga suposiciones relativamente vagas acerca del manual, ideas un tanto más claras acerca de la tecnología, y costos y objetivos de desempeño bien definidos. Puede comenzar diseñando los flujos de datos, las secuencias de control, los conceptos de la densidad del empaque, etc. Idea o adapta las herramientas necesarias, especialmente el sistema de registro, incluyendo el sistema de automatización del diseño.

Mientras tanto, a nivel de la realización, hay que diseñar, mejorar y documentar los circuitos, las tarjetas, los cables, los armazones, las fuentes de poder, y las memorias. Este trabajo prosigue en paralelo con la arquitectura y la implementación.

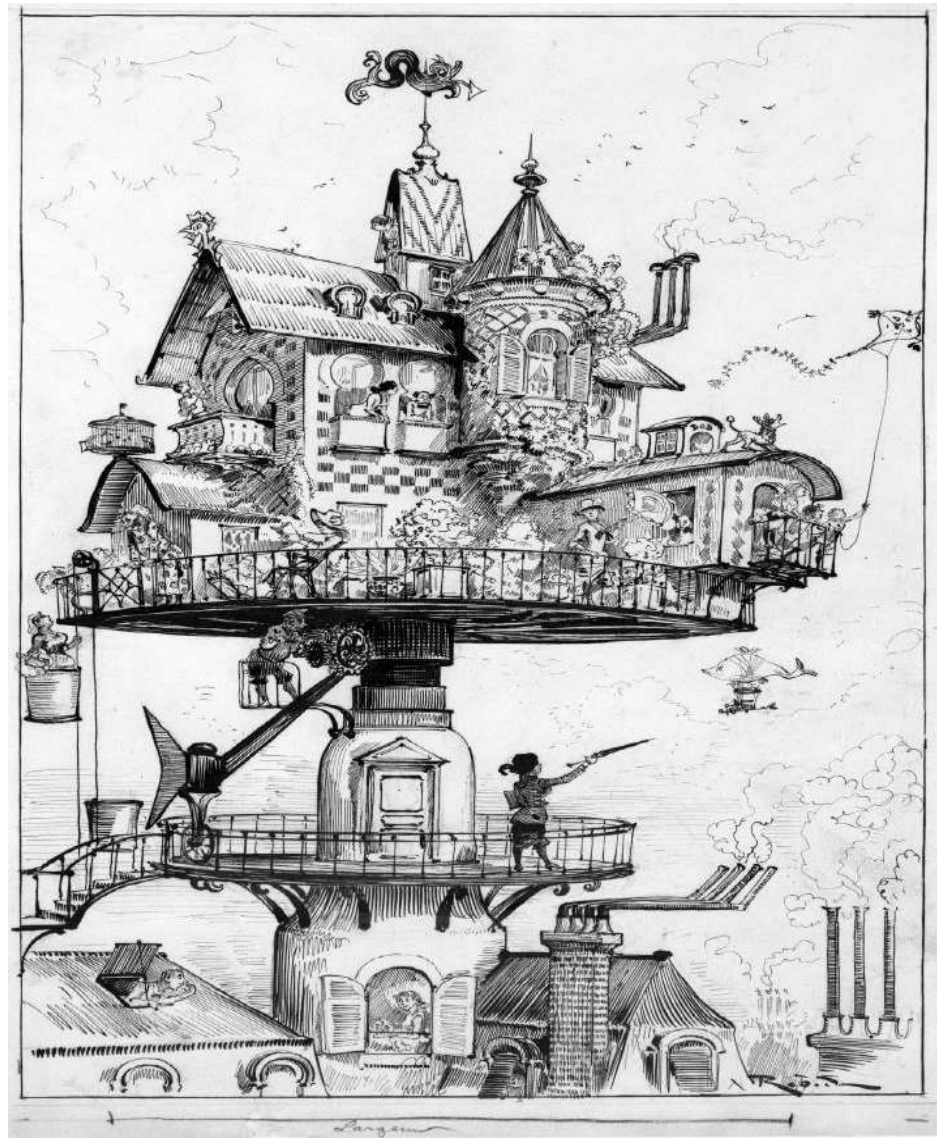
Lo mismo es cierto en el diseño del sistema de programación. Mucho antes de que se completen las especificaciones externas, el implementador tiene mucho trabajo. Puede proceder, dadas algunas vagas aproximaciones en cuanto a las funciones del sistema que finalmente se incorporarán en las especificaciones externas. Debe tener bien definidos los objetivos de espacio y tiempo. Debe conocer la configuración del sistema en el que se ejecutará su producto. Después puede empezar diseñando el alcance de los módulos, las estructuras de tablas, los desgloses de paso o fase, los algoritmos, y todo tipo de herramientas. También, debe invertir algo de tiempo en la comunicación con el arquitecto.

Mientras tanto, en el nivel de la realización también hay mucho que hacer. La programación también tiene una tecnología. Si la máquina es nueva, se debe trabajar mucho en las convenciones de las subrutinas, en las técnicas de supervisión y en los algoritmos de búsqueda y clasificación.⁷

La integridad conceptual requiere que un sistema refleje una sola filosofía y que la especificación vista por el usuario fluya de unas pocas mentes. Sin embargo, debido a la división real del trabajo entre arquitectura, implementación, y realización, esto no implica que un sistema diseñado de esta forma demore más en construirse. La experiencia muestra lo contrario, que todo el sistema junto va más rápido y las pruebas demoran menos. De hecho, una división del trabajo esparcida de forma horizontal ha sido reducida tajantemente a través de una división vertical del trabajo y el resultado son unas comunicaciones simplificadas de forma radical y una integridad conceptual mejorada.

5

El Efecto del Segundo-Sistema



5

El Efecto del Segundo-Sistema

Adde parvum parvo magnus acervus erit.
[Poco a poco se hace un gran montón.]

OVIDIO

Girando la casa por el tráfico. Litografía, Paris, 1882
De *Le Vingtième Siècle* por A. Robida

Si separamos la responsabilidad de la especificación funcional de la responsabilidad de construir un producto rápido y barato, ¿Qué mecanismo limita el entusiasmo inventivo del arquitecto?

La clave esencial es la exhaustiva, cuidadosa, y comprensiva comunicación entre el arquitecto y el constructor. Sin embargo, existen cuestiones de grano más fino que merecen atención.

Procedimiento Interactivo para el Arquitecto

El arquitecto de una construcción trabaja en contra de un presupuesto, usa técnicas de estimación que luego son confirmadas o corregidas por las ofertas del contratista. Con frecuencia sucede que todas las propuestas exceden el presupuesto. Entonces el arquitecto mejora su estimación y reduce su diseño y realiza otra iteración. Quizá le pueda sugerir al contratista maneras de implementar su diseño de forma más barata de la que había ideado.

Un proceso análogo rige al arquitecto de un sistema de computación o de un sistema de programación. Él tiene, además, la ventaja de obtener ofertas del contratista en muchas etapas tempranas del diseño, casi en cualquier momento que la solicite. Aunque, en general, tiene la desventaja de trabajar con un solo contratista, que aumenta o disminuye sus estimaciones para reflejar su acuerdo con el diseño. En la práctica, la comunicación pronta y continua puede brindar al arquitecto buenos indicios del costo y al constructor confianza en el diseño sin disipar la clara división de responsabilidades.

El arquitecto cuando se enfrenta con una estimación muy alta tiene dos opciones posibles: recortar el diseño o enfrentar la estimación sugiriendo implementaciones más baratas. La última es intrínsecamente una tarea emocionante. Ahora el arquitecto está retando al constructor acerca de la forma de hacer su trabajo de construcción. Para que esto tenga éxito, el arquitecto debe

- recordar que el constructor tiene la responsabilidad inventiva y creativa para realizar la implementación; así que el arquitecto sugiere, no impone;
- estar siempre preparado para sugerir *una* manera de implementar cualquier cosa que se especifique y así también para aceptar cualquier otra forma de cumplir con los objetivos;

- tratar de forma tranquila y privada tales sugerencias;
- estar listo a renunciar al crédito por las mejoras sugeridas.

El constructor, en general, será contrario a las sugerencias de cambiar la arquitectura. A menudo estará en lo cierto – alguna característica menor puede tener grandes costos inesperados cuando se elabore la implementación.

Auto-Disciplina – El efecto del Segundo-Sistema

El primer trabajo de un arquitecto es adecuado para ser sobrio y limpio. Él sabe que no sabe qué es lo que está haciendo, así que lo hace cuidadosamente y con grandes restricciones.

En tanto diseña el primer trabajo, se le ocurre adorno tras adorno y embellecimiento tras embellecimiento. Todos esto queda almacenado para ser usado la “siguiente vez”. Tarde o temprano el primer sistema está terminado, y el arquitecto, con firmeza, seguridad y una experiencia demostrada en tal clase de sistemas, está listo para construir el segundo sistema.

Este segundo es el sistema más peligroso que una persona jamás diseña. Cuando haga su tercer y los demás, sus experiencias previas se confirmarán mutuamente como las características generales de dichos sistemas, y sus diferencias identificarán esas partes de su experiencia que son particulares y no generalizables.

La tendencia general es sobrediseñar el segundo sistema, utilizando toda las ideas y adornos que cuidadosamente se hicieron a un lado en el primer sistema. El resultado, como dice Ovidio, es un “gran montón”. Por ejemplo, considere la arquitectura de la IBM 709, después incorporada en la 7090. Ésta es una actualización, un segundo sistema de la muy exitosa y limpia 704. El conjunto de operaciones es tan rico y profuso que solo cerca de la mitad se usaba regularmente.

Veamos un caso más sólido, la arquitectura, la implementación, e incluso la realización de la computadora Stretch, un desahogo para los deseos inventivos reprimidos de mucha gente, y un segundo sistema para la mayoría de ellos. Como afirma Strachey en su reseña:

Tengo la impresión de que el Stretch es de alguna manera el fin de una tendencia de desarrollo. Como algunos primeros programas de computación es inmensamente ingenioso, inmensamente complicado, y extremadamente eficaz, pero de alguna manera al mismo tiempo crudo, excesivo y torpe, y uno siente que debe haber una mejor forma de hacer las cosas.¹

El Operative System/360 fue el segundo sistema para la mayoría de sus diseñadores. Los grupos de diseñadores venían de construir el sistema operativo de disco 7010, el sistema operativo Stretch, el sistema de tiempo real Project Mercury, y el IBSYS para la 7090. Casi nadie tenía experiencia con dos sistemas operativos previos.² Así que el OS/360 es un excelente ejemplo del efecto del segundo-sistema, un Stretch del arte del software al cual se aplican sin cambios tanto el elogio como el reproche de Strachey.

Por ejemplo, el OS/360 dedicaba 26 bytes a una rutina residente de cambio de fecha para el apropiado manejo del 31 de Diciembre en los años bisiestos (cuando es el Día 366). Esto pudo haberse dejado al operador.

El efecto del segundo-sistema tiene otra manifestación algo distinta del adorno funcional puro. Y es una tendencia a refinar técnicas cuya misma existencia se ha vuelto obsoleta por los cambios en los supuestos básicos del sistema. El OS/360 tiene muchos ejemplos de este tipo.

Considere el editor de enlace, diseñado para cargar programas compilados de forma separada y resolver sus referencias-cruzadas. Más allá de esta función básica también maneja las superposiciones del programa. Es uno de los medios de superposición más finos jamás contruidos. Permite estructurar la superposición externamente, en tiempo de enlace, sin necesidad de diseñarlo en el código fuente. Permite cambiar la estructura de la superposición de ejecución a ejecución sin recompilación. Facilita una rica variedad de opciones útiles y otros medios. En cierto sentido es la culminación de años de desarrollo de la técnica de superposición estática.

Sin embargo, es también el último y más fino de los dinosaurios, porque pertenece a un sistema en el que la multiprogramación es el modo normal y la asignación dinámica del núcleo es el supuesto principal. Esto entra en un conflicto directo con la noción de usar superposiciones estáticas. Cuánto mejor funcionaría el sistema si el esfuerzo dedicado al manejo de las superposiciones se hubiera invertido en lograr que la asignación dinámica del nú-

cleo y los medios de referencia-cruzada dinámica fueran realmente rápidos!

Además, el editor de enlace requiere tanto espacio que contiene por sí mismo muchas superposiciones que aun cuando se usa solo para enlazar sin aplicar el manejador de superposiciones, es más lento que la mayoría de los compiladores del sistema. La ironía de esto es que el propósito del enlazador es evitar la recompilación. Es como el patinador cuyo estómago se adelanta a sus pies, la mejora del sistema avanzó hasta que las suposiciones acerca de éste fueron totalmente rebasadas.

El medio de depuración del TESTRAN es otro ejemplo de esta tendencia. Es la culminación de los medios de depuración por lotes, suministra una copia del estado del sistema realmente elegante y capacidades para el volcado de memoria. Usa el concepto de sección de control y una ingeniosa técnica generadora que permite el trazado selectivo y la copia del estado del sistema sin la sobrecarga del intérprete o la recompilación. Los conceptos creativos del Share Operating System³ para la 709 alcanzaron su máximo esplendor.

Mientras tanto, toda la idea de la depuración por lotes sin recompilación se estaba volviendo obsoleta. Los sistemas de computación interactivos, que usan intérpretes o compiladores incrementales, han resultado ser el desafío más importante. Pero incluso en los sistemas por lotes, la aparición de compiladores de rápida-compilación/lenta-ejecución ha hecho de la depuración a nivel código fuente y de la copia del estado del sistema la técnica preferida. Cuánto mejor hubiera resultado el sistema si se hubiera dedicado, cuanto antes y mejor, mayor esfuerzo al TESTRAN, que a la construcción de medios interactivos y de compilación-rápida!

Otro ejemplo más es el planificador, el cual proporciona realmente excelentes medios para la administración de un flujo de tareas por lotes fijo. En sentido estricto, este planificador es el refinado, mejorado, y embellecido segundo sistema sucesor del 1410-1710 Disk Operative System, un sistema por lotes sin multiprogramación excepto para entradas y salidas y dirigido principalmente a aplicaciones comerciales. Como tal, el planificador del OS/360 funciona bien. Aunque casi está totalmente influenciado por las necesidades del OS/360 de entrada de tareas remotas, multiprogramación, y subsistemas interactivos residentes de forma permanente. Y en efecto, el diseño del planificador hace estas cosas difíciles.

¿Cómo evita el arquitecto el efecto del segundo-sistema? Bien, obvia-

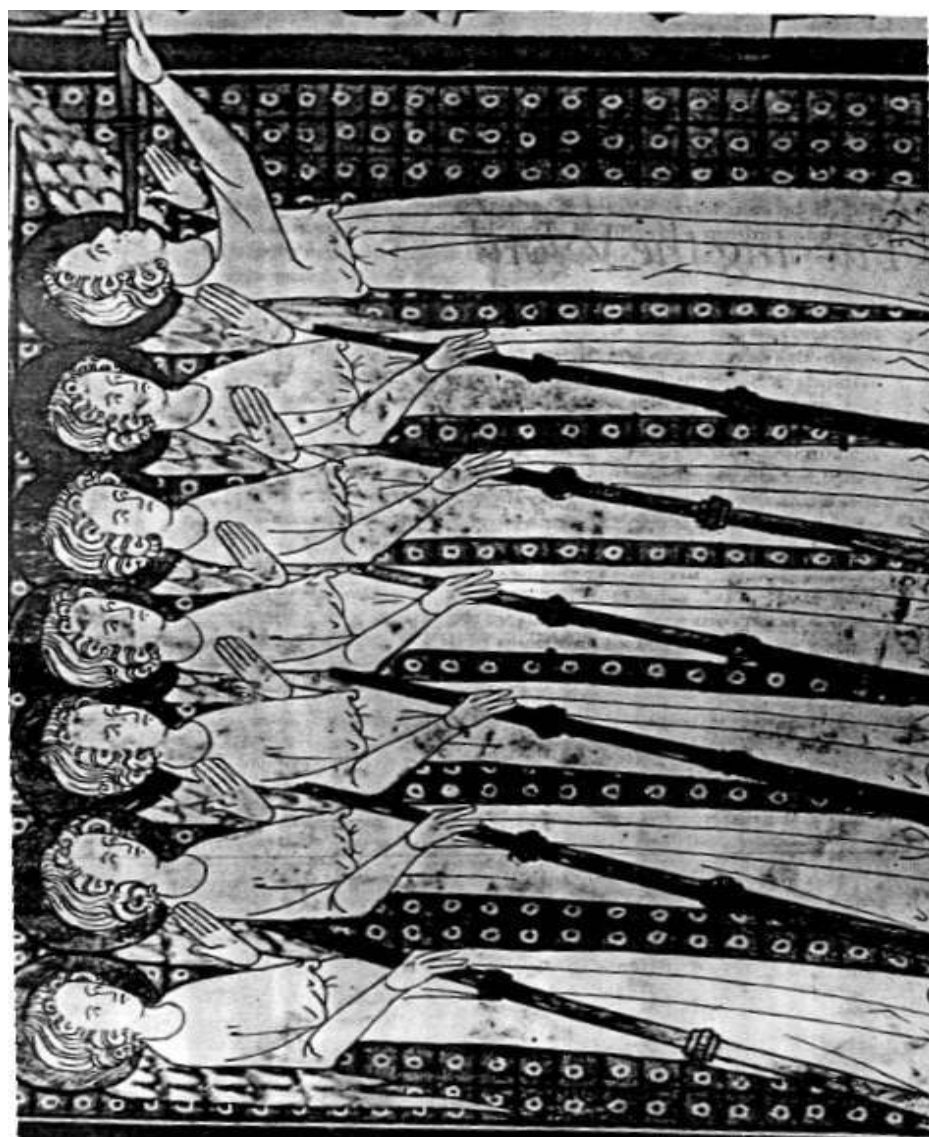
mente no puede evitar su segundo sistema. Pero puede ser consciente de los peligros peculiares de dicho sistema, y ejercer una autodisciplina extra para evitar la ornamentación funcional y la extrapolación de funciones que son innecesarias por cambios en las suposiciones y propósitos.

Un hábito que abrirá los ojos del arquitecto es asignar a cada pequeña función un valor: la habilidad x no vale más que m bytes de memoria y n microsegundos por invocación. Estos valores guiarán las decisiones iniciales y servirán durante la implementación como una guía y alarma para todo.

¿Cómo evita el gestor del proyecto el efecto del segundo-sistema? Exigiendo un arquitecto experimentado que tenga al menos dos sistemas en su haber. Además, permaneciendo atento a las tentaciones particulares, puede hacer las preguntas correctas para así garantizar que los conceptos filosóficos y los objetivos se reflejen totalmente en los detalles del diseño.

6

Pasar la Palabra



6

Pasar la Palabra

Se sentará aquí y dirá, “Haz esto! Haz aquello!” Y nada sucederá.

HARRY S. TRUMAN, ACERCA DEL PODER PRESIDENCIAL ¹

“Las Siete Trompetas” de *The Wells Apocalypse*, siglo 14
El Archivo Bettman

Supongamos que el director tiene un grupo disciplinado y experimentado de arquitectos y muchos implementadores, ¿cómo garantiza que todos escuchen, entiendan e implementen las decisiones tomadas por los arquitectos? ¿Cómo puede un grupo de 10 arquitectos mantener la integridad conceptual de un sistema que está siendo construido por 1000 personas? Para lograr esto se desarrolló toda una tecnología para el diseño de hardware de la System/360, y es igualmente aplicable a proyectos de software.

Especificaciones Escritas – el Manual

El manual, o la especificación escrita, es una herramienta necesaria, aunque no suficiente. El manual es la especificación *externa* del producto. Describe y prescribe cada detalle de lo que el usuario ve. Como tal, es el principal producto del arquitecto.

Su ciclo de preparación es iterativo, conforme la retroalimentación de los usuarios e implementadores muestre dónde el diseño es difícil de usar o construir. Por el bien de los implementadores es importante que los cambios sean cuantizados, esto es, que en el calendario aparezcan versiones fechadas.

El manual no sólo debe describir todo lo que el usuario ve, incluyendo todas las interfaces; sino también debe abstenerse de describir lo que no ve. Esa es la responsabilidad del implementador, y ahí no se debe restringir su libertad de diseño. El arquitecto debe estar siempre preparado para mostrar *una* implementación de cualquier característica que describa, aunque no debe intentar dictar *la* implementación.

El estilo debe ser preciso, completo y detallado con exactitud. Un usuario generalmente se referirá a una sola definición, así que es necesario repetir completamente todo lo indispensable y aún así todo debe ser coherente. Esto ocasiona que los manuales sean de lectura aburrida, pero la precisión es más importante que la vitalidad.

La coherencia del manual System/360's *Principles of Operation*, surge del hecho de que solo fue escrito por dos plumas; la de Gerry Blaauw y la de Andris Padegs. Aunque las ideas son de alrededor de 10 personas, el plasmar esas decisiones en especificaciones en prosa debe ser hecha sólo por una o dos personas, para mantener la coherencia de la prosa y del producto. Para la escritura de una definición se necesitará un montón de minidecisiones que no tienen la importancia de una discusión amplia. Por ejemplo, en la System/360

está el detalle de cómo se coloca el Código de Condición después de cada operación. No es trivial, sin embargo, es el principio de que tales mini-decisiones deben ser coherentes de principio a fin.

Creo que la más fina pieza de escritura de un manual que he visto es el Apéndice de Blaauw del *System/360 Principles of Operation*. Describe con cuidado y precisión los límites de la compatibilidad de la System/360. Define la compatibilidad, prescribe lo que se consigue, y enumera esas áreas del aspecto externo donde la arquitectura se silencia intencionalmente y donde los resultados de un modelo pueden diferir de otros, donde una copia de un modelo dado puede diferir de otra copia, o donde una copia puede diferir incluso de sí misma después de un cambio en la ingeniería. Este es el nivel de precisión al cual aspiran los escritores de manuales, y deben definir lo que *no* está prescrito tan cuidadosamente como lo que *sí* está.

Definiciones Formales

El Inglés o cualquier otro lenguaje humano, por su naturaleza, no es un instrumento de precisión para tales definiciones. Por tal motivo, el escritor del manual debe esforzarse y forzar su lenguaje para conseguir la precisión requerida. Una alternativa interesante es usar una notación formal para tales definiciones. Después de todo, la precisión es la principal cualidad, la *raison d'être* de las notaciones formales.

Vamos a examinar los méritos y las debilidades de las definiciones formales. Como hemos indicado, las definiciones formales son precisas. Tienden a ser completas; los vacíos se hacen más visibles, y así se llenan más rápido. De lo que carecen es de inteligibilidad. Con la prosa en Inglés se pueden mostrar principios estructurales, delinear la estructura en etapas o niveles, y dar ejemplos. Se pueden marcar rápidamente excepciones y destacar contrastes. Y lo más importante, se puede explicar *por qué*. Las definiciones formales presentadas hasta ahora han inspirado maravillas por su elegancia y confianza en su precisión. Aunque han exigido explicaciones en prosa para hacer su contenido más fácil de aprender y enseñar. Por estas razones, creo que las especificaciones futuras constarán de una definición formal y de una definición en prosa.

Un adagio antiguo advierte, “Nunca vayas al mar con dos cronómetros; lleva uno o tres.” Lo mismo se aplica claramente a la prosa y a las definiciones

formales. Si tenemos ambas, el estándar debe ser una, y las demás deben ser las descripciones derivadas, claramente etiquetadas como tal. Cualquiera puede ser el estándar principal. Algol 68 tiene una definición formal como estándar y una definición en prosa como descriptiva. PL/I tiene la prosa como estándar y la descripción formal como derivada. La System/360 tiene también una descripción en prosa como estándar con una descripción formal derivada.

Disponemos de muchas herramientas para las definiciones formales. La Backus-Naur Form es conocida para la definición del lenguaje, y es tratada ampliamente en la literatura.² La descripción formal de PL/I utiliza nuevas ideas de sintaxis abstracta, y está descrita adecuadamente.³ El APL de Iversen ha sido usado para describir máquinas, las más notables la IBM 7090⁴ y la System/360.⁵

Bell y Newell han propuesto nuevas notaciones para describir configuraciones y arquitecturas de máquinas, y las han ilustrado con varias computadoras, entre ellas la DEC PDP-8⁶, la 7090⁶ y la System/360.⁷

Casi todas las definiciones formales surgen para representar o describir una implementación de un sistema de hardware o software cuyas especificaciones externas son descritas por ellas. La sintaxis se puede describir sin esto, pero la semántica generalmente se define al proporcionar un programa dado que realice la operación definida. Esto es, por supuesto, una implementación, y como tal prescribe en exceso la arquitectura. Así pues, se debe tener cuidado al señalar que la definición formal se aplica solo a las especificaciones externas, y se debe decir cuáles son éstas.

Una implementación no sólo es una definición formal, una implementación también se puede usar como una definición formal. Precisamente, esta fue la técnica usada cuando se construyeron las primeras computadoras compatibles. La nueva máquina tenía que ser compatible con una máquina existente. ¿Era vago el manual en algunos puntos? “Pregunte a la máquina!” Se crearía un programa de prueba para determinar el comportamiento, y la nueva máquina se construiría para hacerla compatible.

Un simulador de un sistema de hardware o software puede servir exactamente en la misma forma. Es una implementación; funciona. Así que todas las preguntas de la definición se pueden resolver probándolas.

Usar una implementación como una definición tiene algunas ventajas. Se pueden aclarar todas las preguntas sin ambigüedades a través de experi-

mentos. El debate nunca es necesario, por lo que las respuestas son rápidas. Las respuestas son siempre tan precisas como se quiera, y siempre son correctas, por definición. En oposición a estos se tiene un conjunto formidable de desventajas. La implementación puede prescribir en exceso incluso las especificaciones externas. La sintaxis inválida produce siempre un resultado; en un sistema controlado ese resultado es una indicación de invalidez y *nada más*. En un sistema no controlado pueden aparecer todo tipo de efectos secundarios, y estos pudieron haber sido usados por los programadores. Por ejemplo, cuando emprendimos la simulación de la IBM 1401 sobre la System/360, resultó que habían 30 “curiosidades” – efectos secundarios de supuestas operaciones inválidas – que se habían usado ampliamente y tuvieron que ser consideradas como parte de la definición. La implementación como una definición sobre prescrita; no sólo diría lo que la máquina debe hacer, sino también diría muchas cosas acerca de cómo se tiene que hacer.

Así, también, la implementación a veces dará respuestas inesperadas o no planificadas cuando se hagan preguntas ingeniosas, y resultará que la definición *de facto* generalmente será poco elegante en estas peculiaridades precisamente porque nunca recibieron atención alguna. Esta falta de elegancia generalmente resultará ser lenta o costosa de duplicar en otra implementación. Por ejemplo, algunas computadoras dejan basura en el registro del multiplicando después de una multiplicación. La naturaleza precisa de esta basura resulta ser parte de la definición *de facto*, además duplicándola puede impedir el uso de un algoritmo de multiplicación más veloz.

Finalmente, el uso de una implementación como definición formal es particularmente susceptible a confusión con respecto a si la descripción en prosa o la descripción formal es de hecho el estándar. Esto se cumple especialmente en las simulaciones. También, hay que abstenerse de hacer modificaciones a la implementación mientras se use como estándar.

Incorporación Directa

Una técnica encantadora para diseminar y reforzar las definiciones está a disposición de los arquitectos de sistemas de software. Es especialmente útil para establecer la sintaxis, si no es que la semántica, de las interfaces entre módulos. Esta técnica es para diseñar la declaración de los parámetros asignados o el almacenamiento compartido, y requiere que las implementaciones in-

cluyan esa declaración a través de una operación en tiempo de compilación (un macro o un %INCLUDE en PL/I). Si, además, toda la interfaz está referenciada solo por medio de nombres simbólicos, se puede cambiar la declaración añadiendo o insertando nuevas variables únicamente recompilando, sin alterar el programa en uso.

Reuniones de trabajo y Tribunales

No hace falta decir que las reuniones son necesarias. Las cientos de consultas de persona a persona, deben complementarse con reuniones más largas y formales. Encontramos que dos niveles de éstas resultaron ser útiles. La primera es una reunión de trabajo semanal de medio día de todos los arquitectos, más los representantes oficiales de los implementadores de hardware y software, y los planificadores de mercado. Lo preside el jefe de arquitectos del sistema.

Todos pueden proponer problemas o cambios, aunque las propuestas comúnmente se distribuyen por escrito antes de la reunión. Un nuevo problema normalmente se discute un rato. Se hace énfasis en la creatividad, en lugar de las decisiones únicamente. El grupo intenta idear muchas soluciones a los problemas, luego se pasan unas cuantas soluciones a uno o más arquitectos para su detalle en un manual de propuestas de cambio redactado meticulosamente.

Luego, los detalles de las propuestas de cambio surgen por decisión. Se difunden y son revisados cuidadosamente por implementadores y usuarios y se esbozan con claridad los pros y los contras. Si surge un consenso, tanto mejor. Si no, el jefe de arquitectos decide. Se conservan las actas y las decisiones se difunden de manera formal, puntual, y amplia.

Las decisiones que resultan de las reuniones semanales dan resultados rápidos y permiten que el trabajo avance. Si alguien no está *muy* satisfecho, puede apelar inmediatamente ante el director de proyecto, aunque esto es muy raro que suceda.

Los frutos de estas reuniones brotan de varias fuentes:

1. El mismo grupo – de arquitectos, usuarios, e implementadores – se reúne semanalmente durante meses. No se requiere tiempo para actualizarlos.
2. El grupo es brillante, capaz, bien versado en estas cuestiones, e involucrado profundamente con los resultados. Nadie tiene un papel de “a-

- sesor”. Todos están autorizados para hacer acuerdos vinculantes.
3. Cuando surgen problemas, se buscan soluciones tanto dentro como fuera de las fronteras obvias.
 4. La formalidad de las propuestas escritas concentra la atención, obliga a la toma de decisiones, y evita inconsistencias en el documento elaborado por el comité.
 5. La clara concesión de poder de decisión al jefe de arquitectos evita concesiones y demoras.

A medida que pasa el tiempo, algunas decisiones no se asumen. Algunos asuntos menores no han sido aceptados con entusiasmo por uno u otro participante. Otras decisiones han desarrollado problemas imprevistos, y a veces las reuniones semanales no estuvieron de acuerdo en reconsiderarlas. Así, acumulamos una lista de reclamos menores, cuestiones no resueltas, o disgustos. Para resolver esto, celebramos sesiones anuales del tribunal supremo, que duraban normalmente dos semanas. (Me gustaría mantenerlos cada seis meses si lo hiciera nuevamente.)

Estas sesiones se celebraban un poco antes de las principales fechas de congelamiento del manual. Entre los presentes se incluían no solo a los arquitectos y a los representantes de los programadores e implementadores de la arquitectura, sino también a los gestores de programación, a los de mercadotecnia, y a los implementadores. El director de proyecto de la Sytem/360 presidía. La agenda constaba normalmente de alrededor de 200 temas, en su mayoría de menor importancia, que se enlistaban en tablas y se ubicaban alrededor del salón. Se escuchaban a todos los bandos y se decidía. Por el milagro de la edición de texto computarizada (y mucho por el excelente trabajo del personal), todos los participantes encontraban un manual actualizado cada mañana en su asiento, que plasmaba las decisiones tomadas el día anterior.

Estos “festivales de Otoño” eran útiles no solo para la toma de decisiones, sino también para que fueran aceptadas. Todos eran escuchados, todos participaban, todos entendían mejor las intrincadas restricciones e interrelaciones entre las decisiones.

Implementaciones Múltiples

Los arquitectos de la Sytem/360 tenían dos ventajas casi sin precedentes: suficiente tiempo para trabajar cuidadosamente, e igual influencia política

que los implementadores. El suministro de tiempo suficiente provino del calendario de la nueva tecnología; la igualdad política provino de la construcción simultánea de implementaciones múltiples. La necesidad de una compatibilidad estricta entre estos sirvió como el mejor agente impositivo posible para las especificaciones.

En la mayoría de los proyectos llegaba el día en que se descubría que la máquina y el manual no estaban de acuerdo. Cuando la confrontación continuaba, generalmente el manual perdía, porque éste se podía cambiar de forma más rápida y barata que la computadora. Sin embargo, no tanto cuando existían múltiples implementaciones. Entonces las demoras y los costos asociados con la reparación de la computadora descarriada podían ser superados por las demoras y costos en la revisión de las computadoras que siguieron fielmente el manual.

Podemos aplicar esta misma idea de forma fructífera siempre que estemos definiendo un lenguaje de programación. Uno puede estar seguro que tarde o temprano se tendrán que construir varios intérpretes o compiladores para cumplir diversos objetivos. La definición será más limpia y la disciplina más estricta si desde un inicio se construyen dos implementaciones al menos.

La Bitácora Telefónica

A medida que avanza la implementación, surgen incontables preguntas de interpretación arquitectónica, sin importar cuán precisa sea la especificación. Sin duda, muchas de tales preguntas necesitan ampliaciones y aclaraciones en el texto. Otras solo reflejan malentendidos.

Sin embargo, es importante animar al implementador confundido a telefonar al arquitecto responsable y preguntarle, en lugar de adivinar y continuar. Es igualmente vital reconocer que las respuestas a tales preguntas son declaraciones arquitectónicas *ex cathedra* que deben ser comunicadas a todos.

Un mecanismo útil es una *bitácora telefónica* mantenida por el arquitecto. En ella registra cada pregunta y cada respuesta. Cada semana las bitácoras de varios arquitectos se concatenan, se reproducen, y se distribuyen a los usuarios e implementadores. Aunque este mecanismo es bastante informal, es rápido y además exhaustivo.

La prueba del Producto

El mejor aliado de un director de proyecto es su adversario cotidiano, y es la formación de un probador del producto independiente. Este grupo comprueba las computadoras y los programas en contra de las especificaciones y sirve como abogado del diablo, señala cada defecto y discrepancia concebible. Toda organización de desarrollo necesita tal grupo de auditoría técnica independiente para mantener su credibilidad.

Al final, el cliente es el auditor independiente. A la luz del implacable uso en el mundo real, cada defecto se hará visible. El grupo de prueba del producto es por tanto el sucedáneo del cliente, está especializado en encontrar defectos. Muchas veces, la minuciosa prueba del producto encontrará lugares donde la palabra no aprobó, donde las decisiones de diseño no fueron entendidas apropiadamente o implementadas exactamente. Por este motivo, dicho grupo de prueba es un eslabón necesario en la cadena a través de la cual se aprueba la palabra de diseño, un eslabón que debe operar pronta y simultáneamente con el diseño.

7

*¿Por Qué Fracasó la
Torre de Babel?*



7

¿Por Qué Fracasó la la Torre de Babel?

Ahora bien, toda la tierra continuaba siendo de un solo lenguaje y de un solo conjunto de palabras. Y aconteció que, al ir viajando hacia el este, finalmente descubrieron una llanura-valle en la tierra de Sinar, y se pusieron a morar allí. Y empezaron a decirse, cada uno al otro: ¡Vamos! Hagamos ladrillos y cozámoslos con un procedimiento de quema. De modo que el ladrillo les sirvió de piedra, pero el betún les sirvió de argamasa. Entonces dijeron: ¡Vamos! Edifiquémonos una ciudad y también una torre con su cúspide en los cielos, y hagámonos un nombre célebre, por temor de que seamos esparcidos por toda la superficie de la tierra. Y el señor procedió a bajar para ver la ciudad y la torre que los hijos de los hombres habían edificado. A continuación dijo el señor: ¡Mira! Son un solo pueblo y hay un solo lenguaje para todos ellos, y esto es lo que comienzan a hacer. Pues, ahora no hay nada que tengan pensado hacer que no les sea posible lograr. ¡Vamos! Bajemos y confundamos allí su lenguaje para que no escuche el uno el lenguaje del otro. Por consiguiente, el señor los esparció desde allí sobre toda la superficie de la tierra, y poco a poco dejaron de edificar la ciudad.

GENESIS 11:1-8

“Turmbau zu Babel,” P. Breughel, el Viejo, 1563
Museo Kunsthistorisches, Viena

Una Auditoría de Gestión del Proyecto Babel

De acuerdo al relato del Génesis, la torre de Babel fue el segundo gran emprendimiento de ingeniería del hombre, después del arca de Noé. Babel fue el primer fiasco de ingeniería.

La leyenda es profunda e instructiva en varios niveles. Sin embargo, examinémoslo simplemente como un proyecto de ingeniería, y observemos qué lecciones de gestión podemos aprender. ¿Qué tan bien estaba equipado su proyecto con los prerequisites para el éxito? ¿Los tenían?

1. ¿Una *misión clara*? Sí, aunque ingenuamente imposible. El proyecto fracasó mucho antes de toparse con esta limitación fundamental.
2. ¿*Mano de obra*? Abundante
3. ¿*Materiales*? La arcilla y el asfalto son abundantes en Mesopotamia
4. ¿*Tiempo* suficiente? Sí, no existe indicio de ninguna restricción de tiempo
5. ¿*Tecnología* adecuada? Sí, la estructura piramidal o cónica es intrínsecamente estable y distribuye bien las cargas de compresión. Claramente la albañilería estaba bien entendida. El proyecto fracasó antes de chocar con las limitaciones tecnológicas.

Ahora bien, si tenían todas estas cosas, ¿por qué fracasó el proyecto? ¿Dónde estuvieron sus carencias? En dos aspectos – *la comunicación*, y su consecuencia, *la organización*. Fueron incapaces de hablar entre ellos; por lo tanto no pudieron coordinarse. Cuando falló la coordinación, el trabajo se detuvo. Leyendo entre líneas deducimos que la falta de comunicación condujo a disputas, malos sentimientos, y celos de grupo. En resumen, los clanes empezaron a apartarse y prefirieron el aislamiento a la disputa.

La Comunicación en los Proyectos de Programación Grandes

Así es actualmente. El desastre en la planificación, los desajustes funcionales, y los errores del sistema todos surgen porque la mano izquierda no sabe lo que hace la derecha. Conforme avanza el trabajo, los distintos grupos cambian lentamente las funciones, los tamaños, las velocidades de sus propios programas, y cambian explícita o implícitamente sus suposiciones acerca de las entradas disponibles y los usos que se harán de las salidas.

Por ejemplo, el implementador de una función de superposición de programas puede toparse con problemas y reducir la *velocidad* basándose en estadísticas que muestran cuán rara vez aparecerá esta función en los programas de aplicación. Mientras tanto, de regreso a casa, sus vecinos podrían estar diseñando una parte importante del supervisor que depende de manera crítica de la velocidad de esta función. En sí, este cambio en la velocidad viene a ser un gran cambio en la especificación, y debe ser comunicado ampliamente y ponderado desde el punto de vista del sistema.

¿Cómo, entonces, se debe comunicar un equipo con otro? De tantas formas como sea posible.

- *Informalmente.* Un buen servicio telefónico y una clara definición de las dependencias intergrupales fomentarán las cientos de llamadas de las que depende la interpretación común de los documentos escritos.
- *Reuniones.* Son invaluable las reuniones habituales del proyecto, con uno y otro equipo, que brinden información técnica. De esta forma se disipan cientos de malentendidos menores.
- *Libro de Trabajo.* Se debe empezar un libro de trabajo del proyecto oficial desde un principio. Esto merece una sección aparte.

El Libro de Trabajo del Proyecto

Qué es. El libro de trabajo del proyecto no es tanto un documento separado sino una estructura impuesta a los documentos que el proyecto producirá de todos modos.

Todos los documentos del proyecto deben ser parte de esta estructura. Esto incluye objetivos, especificaciones externas, especificaciones de interfaces, estándares técnicos, especificaciones internas, y memorándums administrativos.

Por qué. La prosa técnica es casi inmortal. Si se examina la genealogía de un manual de usuario de una pieza de hardware o software, se pueden rastrear no sólo las ideas, sino también muchas de las oraciones y los párrafos precisos al primer memorándum proponiendo el producto o explicando el primer diseño. Para el escritor de documentos técnicos, el pegamento es tan poderoso como la pluma.

Ya que esto es así, y puesto que los manuales de la calidad del producto de mañana crecerán a partir de los memorándums de hoy, es muy importante obtener la estructura correcta de la documentación. El pronto diseño del libro de trabajo del proyecto nos garantiza que la propia estructura de la documentación se hizo de forma organizada, no descuidada. Además, la creación de una estructura da forma a la inclusión posterior de segmentos que se ajustan a dicha estructura.

El segundo motivo para el libro de trabajo es el control de la distribución de la información. El problema no es restringir la información, sino asegurar que la información relevante llegue a todos los que la necesiten.

El primer paso es numerar todos los memorándums, de tal manera que se dispongan de listas ordenadas de títulos y cada trabajador pueda ver si encuentra lo que busca. La organización del libro de trabajo va más allá de esto al establecer una estructura de árbol de memorándums. La estructura de árbol permite que las listas de distribución sean mantenidas por los subárboles, si así se desea.

Procedimiento. Como con muchos problemas de gestión de la programación, el problema del memorándum técnico empeora de forma no lineal a medida que aumenta el tamaño. Con 10 personas, los documentos simplemente se pueden numerar. Con 100 personas, generalmente bastarán varias secuencias lineales. Con 1000 personas, se dispersan inevitablemente sobre varios espacios físicos, aumenta la *urgencia* por un libro de trabajo organizado y también aumenta el *tamaño* del libro de trabajo. ¿Entonces, cómo se debe manejar este procedimiento?

Pienso que esto se hizo bien en el proyecto del OS/360. O. S. Locken insistió encarecidamente por un libro de trabajo bien organizado, él había visto su eficacia en su proyecto previo, el sistema operativo 1410-7010.

Rápidamente decidimos que *cada* programador debería ver *todo* el material, i.e., debería tener una copia del libro de trabajo en su propia oficina.

La actualización oportuna tiene una importancia crítica. El libro de trabajo debe estar al corriente. Esto es muy difícil de hacer ya que debido a los cambios constantes todos los documentos se deben volver a escribir. Sin embargo, en un libro de hojas sueltas solo es necesario cambiar páginas. Disponíamos de un sistema de edición de texto por computadora, y demostró ser invaluable para el mantenimiento puntual. Las impresiones principales se pre-

paraban directamente en la impresora de la computadora, y el tiempo de respuesta fue de menos de un día. Sin embargo, el destinatario de todas estas páginas actualizadas tiene un problema de asimilación. Cuando primero recibe una página cambiada, quiere saber, “¿Qué se ha cambiado?” Cuando lo consulta más tarde, quiere saber, “¿Cuál es la definición el día de hoy?”

Esta última exigencia se cumple por el mantenimiento continuo del documento. Resaltar los cambios requiere otros pasos. Primero, se debe marcar en la página el texto modificado, e.g., a través de una barra vertical en el margen a lo largo de cada línea alterada. Segundo, se debe distribuir junto con las nuevas páginas una síntesis, un resumen de cambios escrito de forma separada que enumere los cambios y observaciones acerca de su importancia.

Nuestro proyecto no había andado ni seis meses antes de enfrentarnos a otro problema. El libro de trabajo era de un grosor de cinco pies! Si hubiéramos apilado las 100 copias de servicio de los programadores en nuestras oficina del edificio del Time-Life en Manhattan, se hubieran elevado encima del edificio mismo. Además, la distribución diaria de cambios promediaba 2 pulgadas, algo como 150 páginas a ser intercaladas en todo el documento. El mantenimiento del libro de trabajo empezó a tomar un tiempo significativo en el trabajo diario.

En este punto cambiamos a microfichas, un cambio que ahorró un millón de dólares, incluso nos permitió costear un lector de microfichas para cada oficina. Fuimos capaces de organizar un excelente cambio en la producción de microfichas; el libro de trabajo se encogió de tres pies cúbicos a un sexto de pie cúbico y, mejor aún, las actualizaciones aparecieron en paquetes de cien páginas, reduciendo cien veces el problema del intercalado.

La microficha tiene sus inconvenientes. Desde el punto de vista del gestor el inconveniente de intercalar las páginas aseguraba que los cambios fueran *leídos*, lo cual era el propósito del libro de trabajo. Las microfichas facilitarían bastante el mantenimiento del libro de trabajo, a condición de que las fichas de actualización se distribuyan con un documento en papel que enumere los cambios.

Además, en una microficha, el lector no puede resaltar, marcar, y comentar rápidamente. Los documentos con los cuales el lector ha interactuado son más eficaces para el autor y más útiles para el lector.

Pensándolo bien pienso que la microficha fue un mecanismo muy opor-

tuno, y para proyectos muy grandes lo recomendaría en lugar de un libro de trabajo de papel.

¿Cómo se haría actualmente? Con la actual tecnología de sistemas disponible, creo que la técnica adecuada es el mantenimiento del libro de trabajo en un archivo de acceso directo, marcado con barras de cambio y fechas de revisión. Cada usuario lo consultaría desde una terminal (las máquinas de escribir son muy lentas). Se almacenaría un resumen de cambios, preparado diariamente, en un punto de acceso tipo LIFO. Probablemente el programador lo leería diariamente, pero si falla un día solo deberá leer más tiempo el siguiente día. Mientras lee el resumen de cambios, puede interrumpir su lectura y consultar el propio texto cambiado.

Note que el propio libro de trabajo no ha cambiado. Todavía es la colección de documentos de todo el proyecto, estructurado de acuerdo a un diseño prolijo. El único cambio está en los procedimientos de distribución y consulta. D. C. Engelbart y sus colegas en el Stanford Research Institute han construido dicho sistema y lo usan para construir y mantener la documentación de la red ARPA.

D.L. Parnas de la Universidad Carnegie-Mellon ha propuesto una solución aún más radical.¹ Su tesis es que el programador es más eficaz si se lo resguarda de, en lugar de exponerlo a los detalles de construcción de las piezas del sistema que no sean suyos. Esto presupone que todas las interfaces están definidas de forma completa y precisa. A pesar de que eso es definitivamente un diseño sano, depender de su perfecta realización es una receta para el desastre. Un buen sistema de información por un lado expone los errores de las interfaces y por otro estimula su corrección.

La Organización en un Proyecto de Programación Grande

Si en un proyecto existen n trabajadores, existen $(n^2 - n)/2$ interfaces a través de las cuales puede haber comunicación, y hay potencialmente 2^n equipos dentro de los cuales debe existir coordinación. El propósito de la organización es reducir la cantidad de comunicación y coordinación necesaria; por lo tanto, la organización es un ataque radical a los problemas de comunicación tratados arriba.

Los medios por los cuales se evita la comunicación son *la división del trabajo* y *la especialización de funciones*. La estructura tipo árbol de las organizaciones refleja la necesidad cada vez menor de una comunicación detallada cuando se aplica la división y la especialización del trabajo.

De hecho, en realidad la organización tipo árbol surge como una estructura de autoridad y responsabilidad. El principio de que nadie puede servir a dos amos dicta que la estructura de la autoridad sea tipo árbol. Aunque la estructura de comunicación no está tan restringida, y el árbol es una aproximación apenas aceptable de la estructura de comunicación, que es una red. Las deficiencias del enfoque tipo árbol dan origen a grupos de empleados, a grupos de trabajo, comités, e incluso a la organización tipo matriz utilizada en muchos laboratorios de ingeniería.

Consideremos una organización de programación tipo árbol, y examinemos los ingredientes que cualquier subárbol debe tener para ser eficaz. Ellos son:

1. una misión
2. un productor
3. un director técnico o arquitecto
4. un calendario
5. una división del trabajo
6. definiciones de interfaz entre las partes

Todo esto es obvio y normal excepto la distinción entre el productor y el director técnico. Consideremos primero los dos roles, y luego su relación entre ellos.

¿Cuál es el papel del productor? Arma el equipo, divide el trabajo, y establece el calendario. Consigue los recursos necesarios y continúa haciéndolo. Esto significa que una gran parte de su papel es la comunicación externa, hacia arriba y a los lados. Establece el patrón de comunicación e informes dentro del equipo. Por último, garantiza que el calendario se cumpla, moviendo los recursos y la organización para responder al cambio de circunstancias.

¿Qué hay del director técnico? Es el que concibe el diseño que será construido, identifica sus partes secundarias, especifica cómo se verá desde fuera, y esboza su estructura interna. Proporciona unidad e integridad conceptual a todo el diseño; de esta manera, actúa como un límite sobre la complejidad del sistema. A medida que surgen problemas técnicos particulares, inventa soluciones o cambia el diseño del sistema cuando es necesario. Es, en la encantadora frase de Al Capp, “nuestro agente infiltrado haciendo el trabajo sucio.” Sus comunicaciones son principalmente dentro del equipo. Su trabajo es casi completamente técnico.

Ahora está claro que las cualidades requeridas para estos dos roles son bastante diferentes. Estas cualidades vienen en muchas combinaciones diferentes; esa combinación especial plasmada en el productor y el director debe regir la relación entre ellos. Las organizaciones deben ser diseñadas alrededor de las personas disponibles; nadie se ajusta a organizaciones puramente teóricas.

Se pueden dar tres tipos de relaciones, y las tres se encuentran en prácticas exitosas.

El productor y el director técnico pueden ser la misma persona.

Esto se puede hacer fácilmente en equipos muy pequeños, quizás de tres a seis programadores. Es muy poco viable en proyectos más grandes, por dos razones. Primero, rara vez se encuentra la persona con una sólida capacidad para la gestión y una sólida capacidad técnica. Los pensadores son raros; los emprendedores son más raros; y los pensadores-emprendedores son los más raros.

Segundo, en un proyecto más grande cada uno de estos papeles es necesariamente más que un trabajo de tiempo completo. Es difícil para el productor delegar suficientes obligaciones para darse algo de tiempo técnico. Y para el director también es imposible hacer eso sin comprometer la integridad conceptual del diseño.

El productor puede ser el jefe, el director su mano derecha.

En este caso, la dificultad es establecer la *autoridad* del director para la toma de decisiones técnicas sin afectar su tiempo que lo pondría en la cadena de mando de la gestión.

Obviamente el productor debe proclamar la autoridad técnica del director, y debe apoyarlo en una proporción extremadamente alta cuando surjan los casos de prueba. Para que esto sea posible, el productor y el director deben tener una visión semejante acerca de la filosofía de los principios técnicos; deben discutir las principales cuestiones técnicas en privado, antes de que realmente lleguen a un acuerdo; además el productor debe tener un profundo respeto por las habilidades técnicas del director.

Menos obvio es que el productor pueda hacer todo tipo de cosas sutiles con los símbolos de estatus (tamaño de la oficina, la alfombra, el mobiliario, las copias al carbón, etc.) para proclamar que el director, aunque fuera de la línea de gestión, es una fuente del poder de decisión.

Esto puede hacerse para trabajar de forma muy eficaz. Desafortunadamente rara vez se intenta. Los gestores de proyecto han hecho mal su trabajo al utilizar a personal con talento técnico que no tiene cualidades sólidas para la gestión.

El director puede ser el jefe, y el productor su mano derecha.

Robert Heinlein, en *El Hombre Que Vendió la Luna*, describe dicho acuerdo en un ejemplo ilustrativo.

Coster enterró su rostro en sus manos, luego alzó la vista. “Lo sabía, sabía lo que se necesitaba hacer – pero cada vez que trataba de enfrentar un problema técnico un imbécil quería que tomara decisiones acerca de los camiones –o teléfonos– o alguna maldita cosa. Lo siento, Sr. Harriman. Pensé que podía hacerlo.”

Harriman dijo gentilmente. “No dejes que esto te deprima, Bob. ¿No has dormido mucho últimamente, verdad? Te diré lo que vamos a hacer– pondremos a cargo a Ferguson. Tomaré ese escritorio por unos cuantos días y construiré un espacio para protegerte contra tales cosas. Quiero que tu cerebro piense en vectores de reacción, en eficiencia del combustible y cálculo de esfuerzos, no en contratar camiones.” Harriman se paró en la puerta, buscó alguien fuera de la oficina, y notó un hombre que parecía ser el encargado principal de la oficina. “Oiga, usted! Venga aquí.”

El hombre pareció mostrarse sorprendido, se levantó, se acercó a la puerta y dijo: “¿Sí?”

“Quiero que aquel escritorio de la esquina y todas las cosas que están encima sean cambiadas a una oficina vacía en este piso, hágalo ya”

Supervisó el cambio de Coster y su otro escritorio en otra oficina, observó que el teléfono en la nueva oficina estaba desconectado, y, como una ocurrencia repentina, mandó a instalar también un sofá. “Instalaremos un proyector, una máquina de dibujo, un estante y otros accesorios esta misma noche” le dijo a Coster. “Hágame una lista de todo lo que necesite- para trabajos de ingeniería” Regresó a la oficina nominal del jefe de ingeniería y se puso contento tratando de descubrir donde estaba parada la organización y lo que estaba mal en ella.

Unas cuatro horas después acompañó a Berkeley a encontrarse con Coster. El ingeniero en jefe estaba dormido en su escritorio, con la cabeza apoyada en los brazos. Harriman se disponía a marcharse y regresar, pero Coster despertó. “Lo siento” dijo, ruborizado, “Debí quedarme dormido.”

“Este es el motivo de porqué te traje un sofá,” dijo Harriman. “Es más comfortable. Bob le presentó a Jock Berkeley. Es tu nuevo esclavo. Usted continuará como ingeniero en jefe y punto, jefe indiscutible. Jock será el Jefe Mayor de Todo lo Demás. Desde ahora no tienes que preocuparte por nada-excepto por el pequeño detalle de construir una nave Lunar.”

Se estrecharon las manos. “Solo una pregunta, Sr. Coster”, Dijo Berkeley seriamente, “pásame todo lo que quiera – usted dirigirá las cuestiones técnicas – pero por el amor de dios grábelo para que yo sepa qué ocurre. Tendrá un interruptor en su escritorio que operará una grabadora hermética en mi escritorio.”

“Perfecto!” Coster ya se ve más joven, pensó Harriman.

“Y si usted desea algo que no sea técnico, no lo haga. Solo mueva el interruptor y silve; y será hecho!” Berkeley dirigió su mirada a Harriman. “El jefe quiere hablar contigo acerca del trabajo importante. Los dejo y me pondré manos a la obra.” Y se marchó.

Harriman se sentó; Coster hizo lo mismo y dijo, “Waw!”

“¿Se siente mejor?”

“Me ha causado buena impresión ese Berkeley.”

“Tanto mejor; desde ahora él es tu hermano gemelo. Deja de preocuparte; Trabaja antes con él. Pensarás que estás viviendo en un hospital bien acondicionado.”²

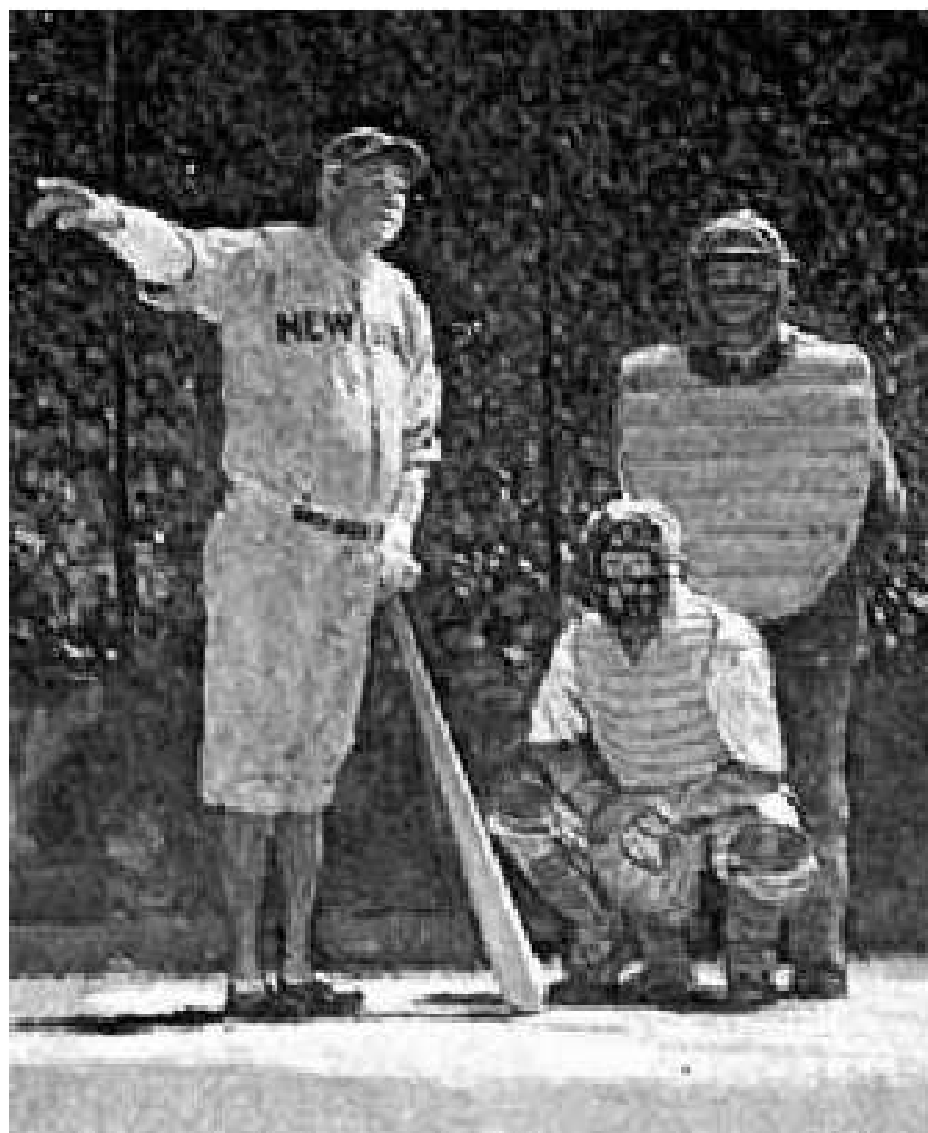
Este relato apenas requiere análisis. Se puede hacer también este arreglo para trabajar de forma eficaz.

Sospecho que el último arreglo es mejor para equipos pequeños, como se discutió en el Capítulo 3, “El Equipo Quirúrgico.” Creo que el arreglo más adecuado es poner al productor como jefe para los subárboles más grandes de un proyecto realmente grande.

La Torre de Babel fue quizá el primer fiasco de ingeniería, pero no fue el último. La comunicación y su consecuencia, la organización, son cruciales para el éxito. Las técnicas de comunicación y organización exigen al gestor mucha reflexión y tanta experiencia como la propia tecnología de software.

8

Predecir el Batazo



8

Predecir el Batazo

La práctica es el mejor de los maestros

PUBLILIUS

La experiencia es un buen maestro, pero los tontos no aprenderán de nadie.

ALMANAQUE DEL POBRE RICHARD

Douglass Crockwell, "Ruth predice su batazo", Serie Mundial, 1932
Reproducido con el permiso de Esquire Magazine y Douglas Crockwell, © 1945
(renovado en 1973) por Esquire, Inc., y la cortesía del Museo Nacional del Béisbol

¿Cuánto tiempo tomará un trabajo de programación de sistemas? ¿Cuánto esfuerzo será necesario? ¿Cómo se estima?

Anteriormente sugerí proporciones que al parecer se aplican a planificar el tiempo asignado a la codificación, a la prueba de componentes, y a la prueba del sistema. Primero, debo decir que *no* se puede estimar la tarea completa sólo estimando la parte de la codificación y luego aplicar las proporciones. La codificación representa sólo alrededor de una sexta parte del problema, y los errores en su estimación o en las proporciones podrían conducir a resultados ridículos.

Segundo, debo decir que los datos para construir programas aislados y pequeños no son aplicables a productos de los sistemas de programación. Por ejemplo, para un programa que en promedio tiene 3200 palabras, el informe de Sackman, Erickson, y Grant revela un tiempo promedio de codificación y depuración de aproximadamente 178 horas para un solo programador, una cifra que si fuera extrapolada daría una productividad anual de 35,800 declaraciones por año. Un programa de la mitad de ese tamaño tomó menos de una cuarta parte de tiempo, y si extrapolamos la productividad obtenemos casi 80,000 declaraciones por año.¹ Hay que añadir el tiempo invertido en la planificación, la documentación, las pruebas, el sistema de integración, y el entrenamiento. La extrapolación lineal de tales cifras de velocidad no tiene sentido. La extrapolación de tiempos para la carrera de 100 yardas muestra que una persona puede correr una milla en menos de tres minutos.

Sin embargo, antes de descartar estas cifras y aunque no son para problemas estrictamente comparables, vemos que sugieren que el trabajo avanza como una potencia del tamaño *incluso* cuando no se involucra ningún tipo de comunicación excepto la de un hombre con sus recuerdos.

La Figura 8.1 nos cuenta la triste historia. Ilustra el resultado del informe de un estudio realizado por Nanus y Farr² en la System Development Corporation. Muestra un exponente de 1.5; eso es,

$$\text{esfuerzo} = (\text{constante}) \times (\text{número de instrucciones})^{1.5}$$

Otro estudio de la SDC reportado por Weinwurm³ también muestra un exponente cercano a 1.5.

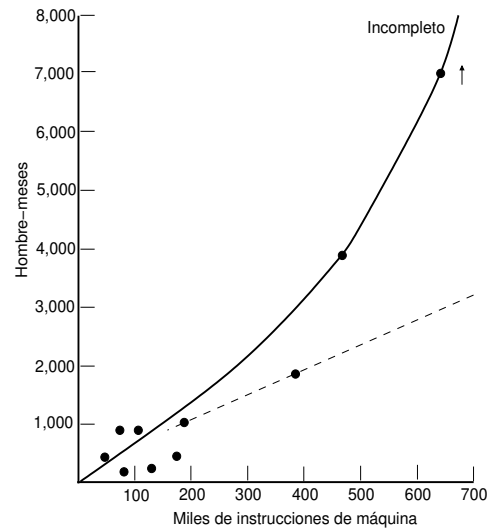


Fig. 8.1 Esfuerzo en función del tamaño del programa

Se han realizado unos cuantos estudios acerca de la productividad de los programadores, y se han propuesto varias técnicas de estimación. Morin ha preparado un estudio de los datos publicados.⁴ Aquí mencionaré sólo unos cuantos puntos que parecen especialmente reveladores.

Los Datos de Portman

Charles Portman, gestor de la División de Software de ICL, de la Computer Equipment Organization (Northwest) en Manchester, ofrece otra útil visión personal.⁵

Encontró que sus equipos de programación erraban en el calendario por cerca de una mitad, pues cada trabajo estaba tomando aproximadamente el doble del tiempo estimado. Las estimaciones fueron realizadas de forma muy prolija por equipos experimentados que estimaron horas-hombre para centenares de subtareas en un diagrama PERT. Cuando apareció el patrón de retraso, les pidió mantener registros minuciosos del uso del tiempo. Estos mostraron que el error en la estimación podía explicarse totalmente por el

hecho de que sus equipos solo estaban realizando el 50 por ciento del trabajo semanal con respecto al tiempo efectivo dedicado a la programación y a la depuración. El resto lo representaba el periodo de inactividad de la máquina, las breves tareas ajenas de mayor prioridad, las reuniones, el papeleo, los negocios de la compañía, la enfermedad, el tiempo dedicado a asuntos personales. En resumen, las estimaciones hicieron una suposición poco realista acerca del número de horas de trabajo técnico por hombre-año. Mi propia experiencia confirma bastante su conclusión.⁶

Los Datos de Aron

Joel Aron, gestor de Systems Technology en IBM en Gaithersburg, Maryland, ha estudiado la productividad del programador cuando trabajaba en nueve sistemas grandes (brevemente, *grande* significa más de 25 programadores y 30,000 instrucciones entregables).⁷ Él divide tales sistemas de acuerdo al número de interacciones entre los programadores (y partes del sistema) y encuentra las siguientes productividades:

Muy pocas interacciones	10,000 instrucciones por hombre-año
Algunas interacciones	5,000
Muchas interacciones	1,500

Los hombre-años no incluyen tareas de apoyo y prueba del sistema, solo diseño y programación. Cuando estas cifras se alteran por un factor de dos para incluir la prueba del sistema, concuerdan mucho con los datos de Harr.

Los Datos de Harr

John Harr, gestor de programación de Electronic Switching System, de la Bell Telephon Laboratories informó acerca de su experiencia y la de otros en un artículo de 1969 en el Spring Joint Computer Conference.⁸ Estos datos se muestran en las figuras 8.2, 8.3, y 8.4.

De éstas, la Fig. 8.2 es la más detallada y útil. Los primeros dos trabajos son básicamente programas de control; los otros dos son básicamente traductores de lenguajes. La productividad se expresa en términos de palabras depuradas por hombre-año. Esto incluye la programación, la prueba de componentes y la prueba del sistema. No está claro cuánto del trabajo de planifica-

	Unids. de prog.	Número de programado- res	Años	Hombre- años	Palabras de programa	Palabras/ hombre- año
Operacional	50	83	4	101	52,000	515
Mantenimiento	36	60	4	81	51,000	630
Compilador	13	9	$2\frac{1}{4}$	17	38,000	2230
Traductor (Ensamblador de Datos)	15	13	$2\frac{1}{2}$	11	25,000	2270

Fig 8.2 Resumen de las cuatro tareas No.1 de programación del ESS

ción, o del trabajo en apoyo de la máquina, la escritura y cosas así, se incluyeron.

Asimismo, las productividades caen en dos clasificaciones; las de los programas de control que son de alrededor de 600 palabras por hombre-año; las de los traductores de cerca de 2200 palabras por hombre-año. Observe que los cuatro programas son de tamaño similar – la variación está en el tamaño de los grupos de trabajo, la duración del tiempo, y número de módulos. ¿Cuál es la causa y cuál es el efecto? ¿Necesitaron los programas de control más personal porque fueron más complicados? ¿O necesitaron más módulos y más hombre-meses porque se les asignó más personal? No se puede estar seguro. Los programas de control seguramente fueron más complejos. Dejando de lado estas incertidumbres, los números describen la productividad real llevada a cabo en un sistema grande, usando las técnicas de programación actuales. Como tales son una contribución auténtica.

Las figuras 8.3 y 8.4 muestran algunos datos interesantes acerca de las tasas de programación y depuración comparadas con las tasas predichas.

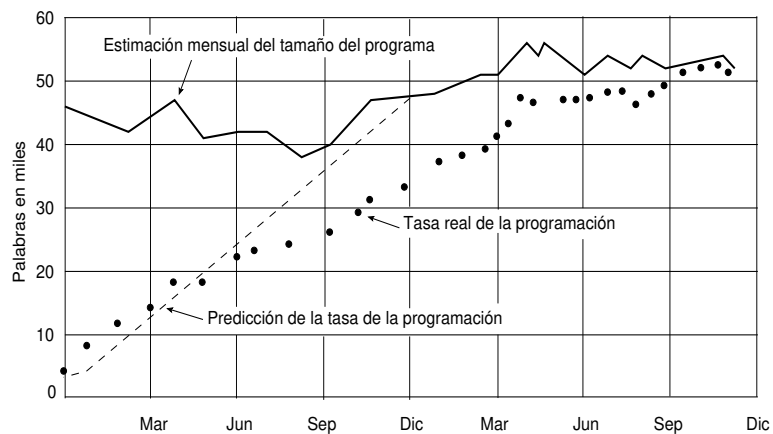


Fig. 8.3 Pronóstico del ESS y tasas de programación reales

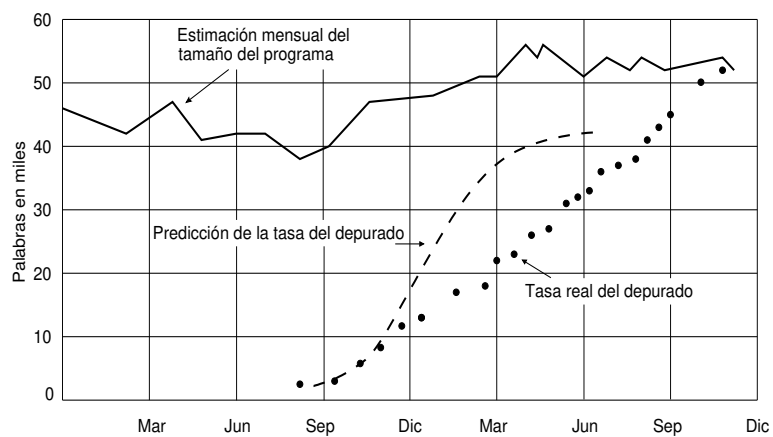


Fig. 8.4 Pronóstico del ESS y tasas de depuración reales

Los Datos del OS/360

La experiencia del IBM OS/360, aunque no está disponible con el detalle de los datos de Harr, lo confirma. Los grupos del programa de control experimentaron productividades en el rango de 600-800 instrucciones depuradas por hombre-año. Los grupos del traductor de lenguajes consiguieron productividades en el rango de 2000-3000 instrucciones depuradas por hombre-año. Esto incluye la planificación realizada por el grupo, la codificación de la prueba del componente, la prueba del sistema, y otras tareas de apoyo. Hasta donde se puede decir, son comparables a los datos de Harr.

Los datos de Aron, los de Harr, y los del OS/360 todos ellos confirman diferencias notables en la productividad relacionada con la complejidad y la dificultad de la tarea misma. Mi pauta en el embrollo de la estimación de la complejidad es que en los compiladores resulta tres veces peor que en los programas normales de aplicaciones por lotes, y en los sistemas operativos la estimación es tres veces peor que en los compiladores.⁹

Los Datos de Corbato

Tanto los datos de Harr como del OS/360 son para la programación en lenguaje ensamblador. Parece que se han publicado pocos datos acerca de la productividad en la programación de sistemas que usan lenguajes de alto nivel. Sin embargo, Corbato, del proyecto MAC del MIT, informa una productividad media de 1200 líneas de declaraciones depuradas en PL/I por hombre-año o en el sistema MULTICS (entre 1 y 2 millones de palabras).¹⁰

Esta cifra es muy interesante. Como los otros proyectos, MULTICS incluye programas de control y traductores de lenguaje. Como los otros, está produciendo un producto de programación de sistemas, probado y documentado. Los datos parecen ser comparables en términos del tipo de trabajo incluido. Y la cifra de productividad es un buen promedio entre el programa de control y las productividades del traductor de otros proyectos.

Aunque la cifra de Corbato es *líneas* por hombre-año, no *palabras*!. Cada declaración en su sistema corresponde a alrededor de cinco palabras de código en ensamblador! Lo que sugiere dos conclusiones importantes.

- La productividad parece constante en términos de declaraciones elementales, una conclusión que es razonable en términos del esfuerzo mental que requiere una declaración y los errores que puede incluir.¹¹
- La productividad de la programación se puede incrementar casi cinco veces cuando se usa un lenguaje de alto nivel adecuado.¹²

9

*Diez Libras
en un Costal de Cinco Libras*



9

Diez Libras en un Costal de Cinco Libras

El autor debería hechar una mirada a Noé, y ... aprender cómo lograron abarrotar en el arca un montón de cosas en un espacio pequeño

SYDNEY SMITH, EDIMBURG REVIEW

**Grabado de una pintura de Heywood Hardy
El archivo Bettman**

El Espacio del Programa como Costo

¿Cuán grande es? Aparte del tiempo de ejecución, el espacio de memoria que ocupa un programa representa un costo importante. Esto es cierto incluso para programas propietario, donde el usuario debe pagar al autor una tarifa que es principalmente una porción del costo de desarrollo. Considere el sistema de software interactivo IBM APL. Se renta por \$400 al mes y, cuando se usa, toma al menos 160 K bytes de memoria. En una Model 165, la memoria se renta a \$12 por kilobytes al mes. Si se dispone del programa a tiempo completo, se paga \$400 por la renta del software y \$1920 por la renta de memoria para usar el programa. Si se usa el sistema APL sólo por cuatro horas al día, los costos son de \$400 por la renta del software y \$320 por la renta de memoria al mes.

Con frecuencia escuchamos expresiones de alarma porque una máquina de 2 M byte pueda tener 400 K dedicados a su sistema operativo. Esto es tan tonto como criticar a un Boeing 747 porque su costo es de \$27 millones. Más bien debemos preguntarnos, “¿Qué es lo que hace?” ¿Qué obtenemos por los dólares invertidos en facilidad de uso y rendimiento (a través del uso eficiente del sistema)? ¿Podrían los \$4800 por mes invertidos en renta de memoria haber sido gastados de manera más provechosa en otro hardware, en programadores o en programas de aplicación?

El diseñador del sistema asigna parte de su recurso total de hardware a memoria para programas residentes cuando piensa que de esta forma hará más por el usuario que como sumadores, discos, etc. Hacerlo de otra manera sería extremadamente irresponsable. Además, el resultado debe ser juzgado como un todo. Nadie puede criticar un sistema de programación por el tamaño *per se* y al mismo tiempo abogar constantemente por una integración más cercana entre el diseño del hardware y el software.

Puesto que el tamaño representa una porción tan grande del costo de uso de un producto de sistemas de programación, el constructor debe establecer metas respecto al tamaño, controlar el tamaño, e idear técnicas para reducir ese tamaño, así como el constructor de hardware establece metas en la cantidad de componentes, controla la cantidad de componentes, e idea técnicas de reducción de componentes. Como cualquier costo, el tamaño como tal no es malo, pero el tamaño innecesario sí lo es.

Control del Tamaño

Para el gestor de proyecto, el control del tamaño es en parte un trabajo técnico y en parte un trabajo de gestión. Se debe estudiar a los usuarios y sus aplicaciones para establecer las características del tamaño de los sistemas a ser promovidos. Luego, estos sistemas deben subdividirse, y asignar a cada componente un determinado tamaño. Puesto que los compromisos en tamaño-velocidad vienen en saltos cuánticos bastante grandes, establecer los objetivos del tamaño es una cuestión delicada, que requiere conocimiento de los compromisos alcanzables dentro de cada pieza. El administrador prudente también debe reservar un fondo, a ser asignado a medida que avanza el trabajo.

En el OS/360, aunque todo esto se hizo de forma prolija, aún tuvimos que aprender otras dolorosas lecciones.

Primero, establecer metas respecto al tamaño del núcleo[†] no es suficiente; se tienen que fijar todos los aspectos relativos al tamaño. En la mayoría de los sistemas operativos anteriores, el sistema residía en cinta, y no se intentaba usarlo descuidadamente para traer segmentos de programa debido a los largos tiempos de búsqueda. El OS/360 residía en disco, como su predecesor inmediato, el Stretch Operating System y el 1410-7019 Disk Operating System. Sus constructores se alegraron por la libertad de los módicos accesos a disco. El resultado inicial fue un rendimiento desastroso.

Cuando asignamos porciones del núcleo a cada componente, no fijamos simultáneamente los costos del acceso. Como esperaba cualquiera que lo viera en retrospectiva, cuando un programador descubría que el programa desbordaba su núcleo asignado entonces lo dividía en superposiciones. Este proceso en sí mismo no solo aumentó el tamaño total sino también ralentizó la ejecución. Y pero aún, nuestro sistema de control de gestión ni media ni se precataba de esto. Todos informaban acerca de la cantidad de *núcleo* que estaban usando, y puesto que estaban dentro de lo establecido, nadie se preocupaba.

Por fortuna, pronto llegó el día en el trabajo en que el simulador de rendimiento del OS/360 empezó a funcionar. El primer resultado indicaba un grave problema. El Fortran H, en una Model 65 con tambores, simulaba la com-

[†]N. del T.: Espacio de memoria asignada al núcleo

pilación a una razón de cinco declaraciones por minuto! La investigación mostró que los módulos del programa de control estaban realizando demasiados accesos a disco. Incluso los módulos supervisores de alta frecuencia estaban haciendo muchos viajes al pozo, y el resultado fue algo bastante parecido a una hiperpaginación.

La primera moraleja es clara: Fijar las cantidades *totales* del espacio de memoria como también del espacio de memoria residente; fijar los accesos al almacenamiento de respaldo así como también los tamaños.

La siguiente lección fue muy similar. La cantidad de espacio se fijó antes de hacer las asignaciones funcionales exactas a cada módulo. El resultado, cualquier programador con problemas de espacio examinaba su código para ver qué podía tirar encima de la cerca al espacio del vecino. Así que la memoria intermedia administrada por el programa de control llegó a ser parte del espacio del usuario. Más grave aún, así se hizo con todo tipo de bloques de control, y el efecto fue comprometer completamente la seguridad y la protección del sistema.

Así pues, la segunda moraleja también es clara: Definir exactamente qué debe hacer un módulo cuando se especifique cuán grande debe ser.

Una tercera y más profunda lección que se aclara a través de estas experiencias. El proyecto era suficientemente grande y la gestión de comunicación suficientemente mala para provocar que muchos miembros del equipo se consideraran concursantes por ganar reconocimiento, en lugar de constructores de productos de programación. Todos suboptimizaban sus piezas por cumplir sus objetivos; unos cuantos se detenían a pensar acerca del efecto total en el cliente. Este falla en la orientación y la comunicación es un gran peligro para proyectos grandes. Durante toda la implementación, los arquitectos del sistema deben mantener una vigilancia continua para asegurar una integridad constante del sistema. Sin embargo, más allá de este mecanismo de vigilancia yace la cuestión de la actitud de los implementadores mismos. Fomentar una actitud de sistema total, y orientado al usuario bien podría ser la función más importante del gestor de programación.

Técnicas de espacio

No hay cantidad de espacio asignado y control que puede hacer pequeño un programa. Esto requiere ingenio y habilidad.

Obviamente, mayor funcionalidad implica mayor espacio, manteniendo la velocidad constante. Así pues, el primer campo de la habilidad está en el compromiso entre funcionalidad y tamaño. Aquí surge una antigua y profunda pregunta de política. ¿Cuánto de dicha elección se debe reservar al usuario? Se puede diseñar un programa con muchos rasgos opcionales, cada uno de los cuales ocupa un poco de espacio. Se puede diseñar un generador que tome una lista de opciones y confeccione un programa para ella. Aunque para cualquier conjunto particular de opciones, un programa más monolítico tomaría menos espacio. Es muy parecido a un auto; si la luz de los mapas, el encendedor, y el reloj se cotizan juntos como una sola opción, el paquete costará menos que si se pudiera escoger cada una de forma separada. Por tanto el diseñador debe decidir cuánto grano fino tendrá la elección de opciones del usuario.

Al diseñar un sistema para un rango de tamaños de memoria, surge otra cuestión primordial. Hay un efecto limitante que evita que el rango apropiado se haga arbitrariamente extenso, incluso con la modularidad de grano fino de la función. En el sistema más pequeño, la mayoría de los módulos estarán superpuestos. Por tanto, es necesario apartar una porción importante del espacio residente del sistema más pequeño como un área temporal o un área de paginado en el que se buscarán otras partes. Su tamaño determina el tamaño de todos los módulos. Además, el dividir funciones en módulos pequeños tiene un costo en espacio y rendimiento. Por lo tanto, un sistema grande que puede permitirse un área temporal veinte veces mayor, de este modo solo ahorrará en accesos. Aún padecerá de velocidad y de espacio porque el tamaño del módulo es muy pequeño. Este efecto limita la máxima eficiencia del sistema que se puede generar a partir de módulos de un sistema pequeño.

El segundo campo de la habilidad son los compromisos espacio-temporales. Para una función dada, mayor espacio implica mayor rapidez. Esto es cierto en un rango impresionantemente extenso. Es este hecho el que hace posible fijar la cantidad de espacio.

El gestor puede hacer dos cosas para ayudar a su equipo a realizar buenos compromisos entre espacio y tiempo. Una es asegurarse de que están capacitados en técnicas de programación, no solo dejarlos depender del ingenio natural y la experiencia previa. Esto es especialmente importante para un

nuevo lenguaje o una nueva máquina. Se debe aprender rápidamente los detalles para usarlos con destreza y compartirlos ampliamente, quizás con premios especiales o elogios por las nuevas técnicas.

La segunda es reconocer que la programación tiene una tecnología, y los componentes necesitan *ser* fabricados. Cada proyecto necesita un cuaderno completo de buenas subrutinas o macros para el manejo de colas, búsqueda, dispersión, y clasificación. Para cada una de dichas funciones la libreta debe tener al menos dos programas, la rápida y la comprimida. El desarrollo de tal tecnología es una tarea importante de entendimiento que se puede realizar en paralelo con la arquitectura del sistema.

La Representación es la Esencia de la Programación

Mas allá de la destreza yace la invención, y es aquí donde nacen los programas ligeros, sobrios, y rápidos. Estos son, casi siempre, el resultado de adelantos estratégicos en lugar de tácticas inteligentes. A veces el adelanto estratégico será un nuevo algoritmo, tal como la Transformada Rápida de Fourier de Cooley-Turkey o la sustitución de una clasificación $n \log(n)$ por un conjunto de n^2 comparaciones.

Es mucho más frecuente que el adelanto estratégico venga de rehacer la representación de los datos o tablas. Aquí es donde yace el corazón de un programa. Enséñeme sus diagramas de flujo y oculte sus tablas, y seguiré desconcertado. Muéstreme sus tablas, y generalmente no necesitaré sus diagramas de flujo; serán obvios.

Es fácil reproducir ejemplos del poder de las representaciones. Recuerdo a un joven emprendiendo la construcción de un complejo intérprete de consola para una IBM 650. Como el espacio era costoso, terminó empaquetándolo en una increíblemente pequeña cantidad de espacio al construir un intérprete para un intérprete, identificando que las interacciones humanas son lentas e infrecuentes. El pequeño y elegante compilador de Fortran de Digitek usa una representación especializada muy densa para el propio código del compilador, por tanto no es necesario el almacenamiento externo. Ese tiempo perdido en la decodificación de esta representación se recupera diez veces al evitar entradas y salidas. (Los ejercicios al final del Capítulo 6 del libro *Automatic Data Processing*¹ de Brooks e Iverson, incluyen una colección de tales

ejemplos, al igual que muchos ejercicios de Knuth.²).

El programador al borde de la desesperación debido a la carencia de espacio a menudo lo mejor que puede hacer es desatenderse de su código, voltear, y contemplar sus datos. La representación es la esencia de la programación.

