ALAN DIX, JANET FINLAY,

GREGORY D. ABOWD, RUSSELL BEALE **HUMAN–COMPUTER INTERACTION**

THIRD EDITION

Human–Computer Interaction

We work with leading authors to develop the strongest educational materials in computing, bringing cutting-edge thinking and best learning practice to a global market.

Under a range of well-known imprints, including Prentice Hall, we craft high quality print and electronic publications which help readers to understand and apply their content, whether studying or at work.

To find out more about the complete range of our publishing, please visit us on the world wide web at: www.pearsoned.co.uk

Human–Computer Interaction Third Edition

Alan Dix, *Lancaster University*

Janet Finlay, *Leeds Metropolitan University*

Gregory D. Abowd, *Georgia Institute of Technology*

Russell Beale, *University of Birmingham*

**Pearson Education Limited**

Edinburgh Gate

Harlow

Essex CM20 2JE

England

and Associated Companies throughout the world

*Visit us on the world wide web at:*

www.pearsoned.co.uk

First published 1993

Second edition published 1998

**Third edition published 2004**

© Prentice-Hall Europe 1993, 1998

© Pearson Education Limited 2004

The rights of Alan Dix, Janet E. Finlay, Gregory D. Abowd and Russell Beale to be identified as authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without either the prior written permission of the publisher or a licence permitting restricted copying in the United Kingdom issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

ISBN-13: 978-0-13-046109-4

ISBN-10: 0-13-046109-1

**British Library Cataloguing-in-Publication Data**

A catalogue record for this book is available from the British Library

10 9 8 7 6 5 4 3

10 09 08 07 06

Typeset in 10/121/2pt Minion by 35

Printed and bound by Scotprint, Haddington

**Part 1**

BRIEF CONTENTS

Guided tour xiv Foreword xvi Preface to the third edition xix Publisher’s acknowledgements xxiii Introduction 1

FOUNDATIONS 9

Chapter 1 The human 11 Chapter 2 The computer 59 Chapter 3 The interaction 123 Chapter 4 Paradigms 164

**Part 2**

DESIGN PROCESS 189

Chapter 5 Interaction design basics 191 Chapter 6 HCI in the software process 225 Chapter 7 Design rules 258 Chapter 8 Implementation support 289 Chapter 9 Evaluation techniques 318 Chapter 10 Universal design 365 Chapter 11 User support 395

**Part 3**

MODELS AND THEORIES 417

Chapter 12 Cognitive models 419 Chapter 13 Socio-organizational issues and stakeholder requirements 450

**vi** Brief Contents

Chapter 14 Communication and collaboratio0n models 475 Chapter 15 Task analysis 510 Chapter 16 Dialog notations and design 544 Chapter 17 Models of the system 594 Chapter 18 Modeling rich interaction 629

**Part 4**

OUTSIDE THE BOX 661

Chapter 19 Groupware 663 Chapter 20 Ubiquitous computing and augmented realities 716 Chapter 21 Hypertext, multimedia and the world wide web 748

References 791 Index 817

**Part 1**

CONTENTS

Guided tour xiv Foreword xvi Preface to the third edition xix Publisher’s acknowledgements xxiii Introduction 1

FOUNDATIONS 9

Chapter 1 The human 11

1.1 Introduction 12 1.2 Input–output channels 13 *Design Focus: Getting noticed* 16 *Design Focus: Where’s the middle?* 22 1.3 Human memory 27 *Design Focus: Cashing in* 30 *Design Focus: 7* ± *2 revisited* 32 1.4 Thinking: reasoning and problem solving 39 *Design Focus: Human error and false memories* 49 1.5 Emotion 51 1.6 Individual differences 52 1.7 Psychology and the design of interactive systems 53 1.8 Summary 55 Exercises 56 Recommended reading 57

Chapter 2 The computer 59

2.1 Introduction 60 2.2 Text entry devices 63 *Design Focus: Numeric keypads* 67 2.3 Positioning, pointing and drawing 71

**viii** Contents

2.4 Display devices 78 *Design Focus: Hermes: a situated display* 86 2.5 Devices for virtual reality and 3D interaction 87 2.6 Physical controls, sensors and special devices 91 *Design Focus: Feeling the road* 94 *Design Focus: Smart-Its – making using sensors easy* 96 2.7 Paper: printing and scanning 97 *Design Focus: Readability of text* 101 2.8 Memory 107 2.9 Processing and networks 114 *Design Focus: The myth of the infinitely fast machine* 116 2.10 Summary 120 Exercises 121 Recommended reading 122

Chapter 3 The interaction 123

3.1 Introduction 124 3.2 Models of interaction 124 *Design Focus: Video recorder* 130 3.3 Frameworks and HCI 130 3.4 Ergonomics 131 *Design Focus: Industrial interfaces* 133 3.5 Interaction styles 136 *Design Focus: Navigation in 3D and 2D* 144 3.6 Elements of the WIMP interface 145 *Design Focus: Learning toolbars* 151 3.7 Interactivity 152 3.8 The context of the interaction 154 *Design Focus: Half the picture?* 155 3.9 Experience, engagement and fun 156 3.10 Summary 160 Exercises 161 Recommended reading 162

Chapter 4 Paradigms 164

4.1 Introduction 165 4.2 Paradigms for interaction 165 4.3 Summary 185

Exercises 186 Recommended reading 187

**Part 2**

Contents **ix**

DESIGN PROCESS 189

Chapter 5 Interaction design basics 191

5.1 Introduction 192 5.2 What is design? 193 5.3 The process of design 195 5.4 User focus 197

*Design Focus: Cultural probes* 200 5.5 Scenarios 201 5.6 Navigation design 203

*Design Focus: Beware the big button trap* 206 *Design Focus: Modes* 207 5.7 Screen design and layout 211 *Design Focus: Alignment and layout matter* 214 *Design Focus: Checking screen colors* 219 5.8 Iteration and prototyping 220 5.9 Summary 222 Exercises 223 Recommended reading 224

Chapter 6 HCI in the software process 225

6.1 Introduction 226 6.2 The software life cycle 226 6.3 Usability engineering 237 6.4 Iterative design and prototyping 241

*Design Focus: Prototyping in practice* 245 6.5 Design rationale 248 6.6 Summary 256

Exercises 257 Recommended reading 257

Chapter 7 Design rules 258

7.1 Introduction 259 7.2 Principles to support usability 260 7.3 Standards 275 7.4 Guidelines 277 7.5 Golden rules and heuristics 282 7.6 HCI patterns 284 7.7 Summary 286

Exercises 287 Recommended reading 288

**x** Contents

Chapter 8 Implementation support 289

8.1 Introduction 290 8.2 Elements of windowing systems 291 8.3 Programming the application 296

*Design Focus: Going with the grain* 301 8.4 Using toolkits 302 *Design Focus: Java and AWT* 304 8.5 User interface management systems 306 8.6 Summary 313 Exercises 314 Recommended reading 316

Chapter 9 Evaluation techniques 318

9.1 What is evaluation? 319 9.2 Goals of evaluation 319 9.3 Evaluation through expert analysis 320 9.4 Evaluation through user participation 327 9.5 Choosing an evaluation method 357 9.6 Summary 362

Exercises 363 Recommended reading 364

Chapter 10 Universal design 365

10.1 Introduction 366 10.2 Universal design principles 366 10.3 Multi-modal interaction 368

*Design Focus: Designing websites for screen readers* 374 *Design Focus: Choosing the right kind of speech* 375 *Design Focus: Apple Newton* 381

10.4 Designing for diversity 384 *Design Focus: Mathematics for the blind* 386 10.5 Summary 393 Exercises 393 Recommended reading 394

Chapter 11 User support 395

11.1 Introduction 396 11.2 Requirements of user support 397 11.3 Approaches to user support 399 11.4 Adaptive help systems 404

*Design Focus: It’s good to talk – help from real people* 405 11.5 Designing user support systems 412 11.6 Summary 414

Exercises 415 Recommended reading 416

**Part 3**

Contents **xi**

MODELS AND THEORIES 417

Chapter 12 Cognitive models 419

12.1 Introduction 420 12.2 Goal and task hierarchies 421 *Design Focus: GOMS saves money* 424 12.3 Linguistic models 430 12.4 The challenge of display-based systems 434 12.5 Physical and device models 436 12.6 Cognitive architectures 443 12.7 Summary 447 Exercises 448 Recommended reading 448

Chapter 13 Socio-organizational issues and stakeholder requirements 450

13.1 Introduction 451 13.2 Organizational issues 451 *Design Focus: Implementing workflow in Lotus Notes* 457 13.3 Capturing requirements 458 *Design Focus: Tomorrow’s hospital – using participatory design* 468 13.4 Summary 472 Exercises 473 Recommended reading 474

Chapter 14 Communication and collaboration models 475

14.1 Introduction 476 14.2 Face-to-face communication 476 *Design Focus: Looking real – Avatar Conference* 481 14.3 Conversation 483 14.4 Text-based communication 495 14.5 Group working 504 14.6 Summary 507 Exercises 508 Recommended reading 509

Chapter 15 Task analysis 510

15.1 Introduction 511 15.2 Differences between task analysis and other techniques 511 15.3 Task decomposition 512 15.4 Knowledge-based analysis 519 15.5 Entity–relationship-based techniques 525 15.6 Sources of information and data collection 532 15.7 Uses of task analysis 538

**xii** Contents

15.8 Summary 541 Exercises 542 Recommended reading 543

Chapter 16 Dialog notations and design 544

16.1 What is dialog? 545 16.2 Dialog design notations 547 16.3 Diagrammatic notations 548

*Design Focus: Using STNs in prototyping* 551 *Design Focus: Digital watch – documentation and analysis* 563 16.4 Textual dialog notations 565 16.5 Dialog semantics 573 16.6 Dialog analysis and design 582 16.7 Summary 589 Exercises 591 Recommended reading 592

Chapter 17 Models of the system 594

17.1 Introduction 595 17.2 Standard formalisms 595 17.3 Interaction models 608 17.4 Continuous behavior 618 17.5 Summary 624

Exercises 625 Recommended reading 627

Chapter 18 Modeling rich interaction 629

18.1 Introduction 630 18.2 Status–event analysis 631 18.3 Rich contexts 639 18.4 Low intention and sensor-based interaction 649

*Design Focus: Designing a car courtesy light* 655 18.5 Summary 657 Exercises 658 Recommended reading 659

**Part 4**

OUTSIDE THE BOX 661

Chapter 19 Groupware 663

19.1 Introduction 664 19.2 Groupware systems 664

Contents **xiii**

19.3 Computer-mediated communication 667 *Design Focus: SMS in action* 673 19.4 Meeting and decision support systems 679 19.5 Shared applications and artifacts 685 19.6 Frameworks for groupware 691 *Design Focus: TOWER – workspace awareness* 701 19.7 Implementing synchronous groupware 702 19.8 Summary 713 Exercises 714 Recommended reading 715

Chapter 20 Ubiquitous computing and augmented realities 716

20.1 Introduction 717 20.2 Ubiquitous computing applications research 717 *Design Focus: Ambient Wood – augmenting the physical* 723 *Design Focus: Classroom 2000/eClass – deploying and evaluating ubicomp* 727 *Design Focus: Shared experience* 732 20.3 Virtual and augmented reality 733 *Design Focus: Applications of augmented reality* 737 20.4 Information and data visualization 738 *Design Focus: Getting the size right* 740 20.5 Summary 745 Exercises 746 Recommended reading 746

Chapter 21 Hypertext, multimedia and the world wide web 748

21.1 Introduction 749 21.2 Understanding hypertext 749 21.3 Finding things 761 21.4 Web technology and issues 768 21.5 Static web content 771 21.6 Dynamic web content 778 21.7 Summary 787

Exercises 788 Recommended reading 788

References 791 Index 817

**xiv** Guided tour

DESIGN PROCESS

In this part, we concentrate on how design practice

addresses the critical feature of an interactive system –

usability from the human perspective. The chapters in this part promote the purposeful design of more usable

interactive systems. We begin in Chapter 5 by introducing

the key elements in the interaction design process. These

elements are then expanded in later chapters.

Chapter 6 discusses the design process in more detail,

specifically focussing on the place of user-centered design

within a software engineering framework. Chapter 7 high

lights the range of design rules that can help us to specify

usable interactive systems, including abstract principles,

guidelines and other design representations.

In Chapter 8, we provide an overview of implementa

tion support for the programmer of an interactive system.

Chapter 9 is concerned with the techniques used to evalu

ate the interactive system to see if it satisfies user needs.

Chapter 10 discusses how to design a system to be univer

sally accessible, regardless of age, gender, cultural background

or ability. In Chapter 11 we discuss the provision of user

support in the form of help systems and documentation.

PART

2 MODELING RICH INTERACTION

OVERVIEW 

We operate within an ecology of people, physical artifacts

and electronic systems, and this rich ecology has recently

become more complex as electronic devices invade the

workplace and our day-to-day lives. We need methods

to deal with these rich interactions.

n Status–event analysis is a semi-formal, easy to apply

technique that:

– classifies phenomena as event or status

– embodies naïve psychology

– highlights feedback problems in interfaces.

n Aspects of rich environments can be incorporated into

methods such as task analysis:

– other people

– information requirements

– triggers for tasks

– modeling artifacts

– placeholders in task sequences.

n New sensor-based systems do not require explicit

interaction; this means:

– new cognitive and interaction models

– new design methods

– new system architectures.

18 

The part structure separates out introductory and more advanced material, with each part opener giving a simple description of what its constituent chapters cover

19.3 Computer-mediated communication **675**

CuSeeMe 

Special-purpose video conferencing is still relatively expensive, but low-fidelity desktop

video conferencing is now within the reach of many users of desktop computers. Digital video

cameras are now inexpensive and easily obtainable. They often come with pre-packaged video

phone or video conferencing software. However, the system which has really popularized

video conferencing is a web-based tool. CuSeeMe works over the internet allowing participants

across the world owning only a basic digital video camera to see and talk to one another. The soft

ware is usually public domain (although there are commercial versions) and the services allowing

connection are often free. The limited bandwidth available over long-distance internet links means

that video quality and frame rates are low and periodic image break-up may occur. In fact, it is

sound break-up which is more problematic. After all, we can talk to one another quite easily with

out seeing one another, but find it very difficult over a noisy phone line. Often participants may

see one another’s video image, but actually discuss using a synchronous text-based ‘talk’ program.



CuSeeMe – video conferencing on the internet. Source: Courtesy of Geoff Ellis

Boxed asides contain descriptions of particular tasks or technologies for additional interest, experimentation and discussion

Bullet points at the start of each chapter highlight the core coverage

**440** Chapter 12 n Cognitive models

Worked exercise *Do a keystroke-level analysis for opening up an application in a visual desktop interface using a mouse as the pointing device, comparing at least two different methods for performing the*

*task. Repeat the exercise using a trackball. Consider how the analysis would differ for various*

*positions of the trackball relative to the keyboard and for other pointing devices.*

Answer We provide a keystroke-level analysis for three different methods for launching an

application on a visual desktop. These methods are analyzed for a conventional one

button mouse, a trackball mounted away from the keyboard and one mounted close to

the keyboard. The main distinction between the two trackballs is that the second one

does not require an explicit repositioning of the hands, that is there is no time required

for homing the hands between the pointing device and the keyboard.

**Method 1** Double clicking on application icon

Steps Operator Mouse Trackball1 Trackball2

1. Move hand to mouse **H**[mouse] 0.400 0.400 0.000

2. Mouse to icon **P**[to icon] 0.664 1.113 1.113

3. Double click **2B**[click] 0.400 0.400 0.400

4. Return to keyboard **H**[kbd] 0.400 0.400 0.000

Total times 1.864 2.313 1.513

**Method 2** Using an accelerator key

Steps Operator Mouse Trackball1 Trackball2

1. Move hand to mouse **H**[mouse] 0.400 0.400 0.000

2. Mouse to icon **P**[to icon] 0.664 1.113 1.113

3. Click to select **B**[click] 0.200 0.200 0.200

4. Pause **M** 1.350 1.350 1.350

5. Return to keyboard **H**[kbd] 0.400 0.400 0.000

6. Press accelerator **K** 0.200 0.200 0.200

Total times 3.214 3.663 2.763

**Method 3** Using a menu

Steps Operator Mouse Trackball1 Trackball2

1. Move hand to mouse **H**[mouse] 0.400 0.400 0.000

2. Mouse to icon **P**[to icon] 0.664 1.113 1.113

3. Click to select **B**[click] 0.200 0.200 0.200

4. Pause **M** 1.350 1.350 1.350

5. Mouse to file menu **P** 0.664 1.113 1.113

6. Pop-up menu **B**[down] 0.100 0.100 0.100

7. Drag to open **P**[drag] 0.713 1.248 1.248

8. Release mouse **B**[up] 0.100 0.100 0.100

9. Return to keyboard **H**[kbd] 0.400 0.400 0.000

Total times 4.591 6.024 5.224

Worked exercises within chapters provide step-by-step guidelines to demonstrate problem-solving techniques

Guided tour **xv**

**732** Chapter 20 n Ubiquitous computing and augmented realities

within these environments. Much of our understanding of work has developed from

Fordist and Taylorist principles on the structuring of activities and tasks. Evaluation

within HCI reflects these roots and is often predicated on notions of task and the

measurement of performance and efficiency in meeting these goals and tasks.

However, it is not clear that these measures can apply universally across activities

when we move away from structured and paid work to other activities. For example,

DESIGN FOCUS 

Shared experience

You are in the Mackintosh Interpretation Centre in an arts center in Glasgow, Scotland. You notice a

man wearing black wandering around looking at the exhibits and then occasionally at a small PDA he is

holding. As you get closer he appears to be talking to himself, but then you realize he is simply talking

into a head-mounted microphone. ‘Some people can never stop using their mobile phone’, you think.

As you are looking at one exhibit, he comes across and suddenly cranes forward to look more closely,

getting right in front of you. ‘How rude’, you think.

The visitor is taking part in the City project – a mixed-reality experience. He is talking to two other

people at remote sites, one who has a desktop VR view of the exhibition and the other just a website.

However, they can all see representations of each other. The visitor is being tracked by ultrasound and

he appears in the VR world. Also, the web user’s current page locates her in a particular part of the

virtual exhibition. All of the users see a map of the exhibitiion showing where they all are.

**10.5**

EXERCISES

You might think that in such an experiment the person actually in the museum would take the lead, but

in fact real groups using this system seemed to have equal roles and really had a sense of shared experi

ence despite their very different means of seeing the exhibition.

See the book website for a full case study: /e3/casestudy/city/



City project: physical presence, VR interfaces and web interface. Source: Courtesy of

Matthew Chalmers, note: City is an Equator project

Design Focus mini case studies highlight practical

applications of HCI concepts

Exercises **393**

SUMMARY

Universal design is about designing systems that are accessible by all users in all

circumstances, taking account of human diversity in disabilities, age and culture.

Universal design helps everyone – for example, designing a system so that it can be

used by someone who is deaf or hard of hearing will benefit other people working in

noisy environments or without audio facilities. Designing to be accessible to screen

reading systems will make websites better for mobile users and older browsers.

Multi-modal systems provide access to system information and functionality

through a range of different input and output channels, exploiting redundancy.

Such systems will enable users with sensory, physical or cognitive impairments to

make use of the channels that they can use most effectively. But all users benefit

from multi-modal systems that utilize more of our senses in an involving interactive

experience.

For any design choice we should ask ourselves whether our decision is excluding

someone and whether there are any potential confusions or misunderstandings in

our choice.



Frequent links to the

book website for

further information

Recommended reading **509**

RECOMMENDED READING

J. Carroll, editor, *HCI Models, Theories, and Frameworks: Toward an Interdisciplinary Science*, Morgan Kaufmann, 2003.

See chapters by Perry on distributed cognition, Monk on common ground and Kraut on social psychology.

L. A. Suchman, *Plans and Situated Actions: The Problem of Human–Machine Communication*, Cambridge University Press, 1987.

This book popularized ethnography within HCI. It puts forward the viewpoint that most actions are not pre-planned, but situated within the context in which they occur. The principal domain of the book is the design of help for a photo copier. This is itself a single-user task, but the methodology applied is based on both ethnographic and conversational analysis. The book includes several chap ters discussing the contextual nature of language and analysis of conversation transcripts.

T. Winograd and F. Flores, *Understanding Computers and Cognition: A New Foundation for Design*, Addison-Wesley, 1986.

Like Suchman, this book emphasizes the contextual nature of language and the weakness of traditional artificial intelligence research. It includes an account of speech act theory as applied to Coordinator. Many people disagree with the authors’ use of speech act theory, but, whether by application or reaction, this

10.1 Is multi-modality always a good thing? Justify your answer.

10.2 What are (i) auditory icons and (ii) earcons? How can they be used to benefit both visually impaired and sighted users?

10.3 Research your country’s legislation relating to accessibility of technology for disabled people. What are the implications of this to your future career in computing?

10.4 Take your university website or another site of your choice and assess it for accessibility using Bobby. How would you recommend improving the site?

10.5 How could systems be made more accessible to older users?

10.6 Interview either (i) a person you know over 65 or (ii) a child you know under 16 about their experience, attitude and expectations of computers. What factors would you take into account if you were designing a website aimed at this person?

10.7 Use the screen reader simulation available at www.webaim.org/simulations/screenreader to experience something of what it is like to access the web using a screen reader. Can you find the answers to the test questions on the site?

Chapter summaries reinforce student learning. Exercises at the end of chapters can be used by teachers or individuals to test understanding

work has been highly influential.

S. Greenberg, editor, *Computer-supported Cooperative Work and Groupware*,

Academic Press, 1991.

The contents of this collection originally made up two special issues of the

*International Journal of Man–Machine Studies*. In addition, the book contains

Greenberg’s extensive annotated bibliography of CSCW, a major entry point for

any research into the field. Updated versions of the bibliography can be obtained

from the Department of Computer Science, University of Calgary, Calgary,

Alberta, Canada.

*Communications of the ACM*, Vol. 34, No. 12, special issue on ‘collaborative com

puting’, December 1991.

Several issues of the journal *Interacting with Computers* from late 1992 through early

1993 have a special emphasis on CSCW.

*Computer-Supported Cooperative Work* is a journal dedicated to CSCW. See also back

issues of the journal *Collaborative Computing*. This ran independently for a while,

but has now merged with *Computer-Supported Cooperative Work*.

See also the recommended reading list for Chapter 19, especially the conference

proceedings.

Annotated further reading encourages readers to research topics in depth

FOREWORD

Human–computer interaction is a difficult endeavor with glorious rewards. Designing interactive computer systems to be effective, efficient, easy, and enjoyable to use is important, so that people and society may realize the benefits of computation based devices. The subtle weave of constraints and their trade-offs – human, machine, algorithmic, task, social, aesthetic, and economic – generates the difficulty. The reward is the creation of digital libraries where scholars can find and turn the pages of virtual medieval manuscripts thousands of miles away; medical instruments that allow a surgical team to conceptualize, locate, and monitor a complex neuro surgical operation; virtual worlds for entertainment and social interaction, respon sive and efficient government services, from online license renewal to the analysis of parliamentary testimony; or smart telephones that know where they are and under stand limited speech. Interaction designers create interaction in virtual worlds and embed interaction in physical worlds.

Human–computer interaction is a specialty in many fields, and is therefore multi disciplinary, but it has an intrinsic relationship as a subfield to computer science. Most interactive computing systems are for some human purpose and interact with humans in human contexts. The notion that computer science is the study of algo rithms has virtue as an attempt to bring foundational rigor, but can lead to ignoring constraints foundational to the design of successful interactive computer systems.

A lesson repeatedly learned in engineering is that a major source of failure is the narrow optimization of a design that does not take sufficient account of contextual factors. Human users and their contexts are major components of the design problem that cannot be wished away simply because they are complex to address. In fact, that largest part of program code in most interactive systems deals with user interaction. Inadequate attention to users and task context not only leads to bad user interfaces, it puts entire systems at risk.

The problem is how to take into account the human and contextual part of a sys tem with anything like the rigor with which other parts of the system are understood and designed – how to go beyond fuzzy platitudes like ‘know the user’ that are true, but do not give a method for doing or a test for having done. This is difficult to do, but inescapable, and, in fact, capable of progress. Over the years, the need to take into account human aspects of technical systems has led to the creation of new fields of study: applied psychology, industrial engineering, ergonomics, human factors,

Foreword **xvii**

man–machine systems. Human–computer interaction is the latest of these, more complex in some ways because of the breadth of user populations and applications, the reach into cognitive and social constraints, and the emphasis on interaction. The experiences with other human-technical disciplines lead to a set of conclusions about how a discipline of human–computer interaction should be organized if it is to be successful.

First, design is where the action is. An effective discipline of human–computer interaction cannot be based largely on ‘usability analysis’, important though that may be. Usability analysis happens too late; there are too few degrees of freedom; and most importantly, it is not generative. Design thrives on understanding constraints, on insight into the design space, and on deep knowledge of the materials of the design, that is, the user, the task, and the machine. The classic landmark designs in human–computer interaction, such as the Xerox Star and the Apple Lisa/Macintosh, were not created from usability analysis (although usability analysis had important roles), but by generative principles for their designs by user interface designers who had control of the design and implementation.

Second, although the notion of ‘user-centered design’ gets much press, we should really be emphasizing ‘task-centered design’. Understanding the purpose and con text of a system is key to allocating functions between people and machines and to designing their interaction. It is only in deciding what a human–machine system should do and the constraints on this goal that the human and technical issues can be resolved. The need for task-centered design brings forward the need for methods of task analysis as a central part of system design.

Third, human–computer interaction needs to be structured to include both analytic and implementation methods together in the same discipline and taught together as part of the core. Practitioners of the discipline who can only evaluate, but not design and build are under a handicap. Builders who cannot reason analytically about the systems they build or who do not understand the human information pro

cessing or social contexts of their designs are under a handicap. Of course, there will be specialists in one or another part of human–computer interaction, but for there to be a successful field, there must be a common core.

Finally, what makes a discipline is a set of methods for doing something. A field must have results that can be taught and used by people other than their originators to do something. Historically, a field naturally evolves from a set of point results to a set of techniques to a set of facts, general abstractions, methods, and theories. In fact, for a field to be cumulative, there must be compaction of knowledge by crunch ing the results down into methods and theories; otherwise the field becomes fad driven and a collection of an almost unteachably large set of weak results. The most useful methods and theories are generative theories: from some task analysis it is possible to compute some insightful property that constrains the design space of a system. In a formula: task analysis, approximation, and calculation. For example, we can predict that if a graphics system cannot update the display faster than 10 times/second then the illusion of animation will begin to break down. This con straint worked backwards has architectural implications for how to guarantee the needed display rate under variable computational load. It can be designed against.

**xviii** Foreword

This textbook, by Alan Dix, Janet Finlay, Gregory Abowd, and Russell Beale, represents how far human–computer interaction has come in developing and organizing technical results for the design and understanding of interactive systems. Remarkably, by the light of their text, it is pretty far, satisfying all the just

enumerated conclusions. This book makes an argument that by now there are many teachable results in human–computer interaction by weight alone! It makes an argu ment that these results form a cumulative discipline by its structure, with sections that organize the results systematically, characterizing human, machine, interaction, and the design process. There are analytic models, but also code implementation

examples. It is no surprise that methods of task analysis play a prominent role in the text as do theories to help in the design of the interaction. Usability evaluation methods are integrated in their proper niche within the larger framework.

In short, the codification of the field of human–computer interaction in this text is now starting to look like other subfields of computer science. Students by studying the text can learn how to understand and build interactive systems. Human–computer interaction as represented by the text fits together with other parts of computer science. Moreover, human–computer interaction as presented is a challenge problem for advancing theory in cognitive science, design, business, or social-technical systems. Given where the field was just a few short years ago, the

creation of this text is a monumental achievement. The way is open to reap the glorious rewards of interactive systems through a markedly less difficult endeavor, both for designer and for user.

Stuart K. Card

Palo Alto Research Center, Palo Alto, California

PREFACE TO THE THIRD EDITION

It is ten years since the first edition of this book was published and much has changed. Ubiquitous computing and rich sensor-filled environments are finding their way out of the laboratory, not just into films and fiction, but also into our workplaces and homes. Now the computer really has broken its bounds of plastic and glass: we live in networked societies where personal computing devices from mobile phones to smart cards fill our pockets, and electronic devices surround us at home and at work. The web too has grown from a largely academic network into the hub of business and everyday lives. As the distinctions between physical and digital, work and leisure start to break down, human–computer interaction is also radically changing.

We have tried to capture some of the excitement of these changes in this revised edition, including issues of physical devices in Chapters 2 and 3, discussion of web interfaces in Chapter 21, ubiquitous computing in Chapters 4 and 20, and new models and paradigms for interaction in these new environments in Chapters 17 and 18. We have reflected aspects of the shift in use of technology from work to leisure in the analysis of user experience in Chapter 3, and in several of the boxed examples and case studies in the text. This new edition of *Human–Computer Interaction* is not just tracking these changes but looking ahead at emerging areas.

However, it is also rooted in strong principles and models that are not dependent on the passing technologies of the day. We are excited both by the challenges of the new and by the established foundations, as it is these foundations that will be the means by which today’s students understand tomorrow’s technology. So we make no apology for continuing the focus of previous editions on the theoretical and con ceptual models that underpin our discipline. As the use of technology has changed, these models have expanded. In particular, the insular individual focus of early work is increasingly giving way to include the social and physical context. This is reflected in the expanded treatment of social and organizational analysis, including ethnography, in Chapter 13, and the analysis of artifacts in the physical environment in Chapter 18.

**xx** Preface to the third edition

STRUCTURE

The structure of the new edition has been completely revised. This in part reflects the growth of the area: ten years ago HCI was as often as not a minority optional sub ject, and the original edition was written to capture the core material for a standard course. Today HCI is much expanded: some areas (like CSCW) are fully fledged dis ciplines in their own right, and HCI is studied from a range of perspectives and at different levels of detail. We have therefore separated basic material suitable for intro ductory courses into the first two parts, including a new chapter on interaction design, which adds new material on scenarios and navigation design and provides an overview suitable for a first course. In addition, we have included a new chapter on universal design, to reflect the growing emphasis on design that is inclusive of all, regardless of ability, age or cultural background. More advanced material focussing on different HCI models and theories is presented in Part 3, with extended cover age of social and contextual models and rich interaction. It is intended that these sections will be suitable for more advanced HCI courses at undergraduate and postgraduate level, as well as for researchers new to the field. Detailed coverage of the particular domains of web applications, ubiquitous computing and CSCW is given in Part 4.

New to this edition is a full color plate section. Images flagged with a camera icon in the text can be found in color in the plate section.

WEBSITE AND SUPPORT MATERIALS

We have always believed that support materials are an essential part of a textbook of this kind. These are designed to supplement and enhance the printed book – phys ical and digital integration in practice. Since the first edition we have had exercises, mini-case studies and presentation slides for all chapters available electronically. For the second edition these were incorporated into a website including links and an online search facility that acts as an exhaustive index to the book and mini encyclopedia of HCI. For visually disabled readers, access to a full online electronic text has also been available. The website is continuing to develop, and for the third edition provides all these features plus more, including WAP search, multi-choice questions, and extended case study material (see also color plate section). We will use the book website to bring you new exercises, information and other things, so do visit us at www.hcibook.com (also available via www.booksites.net/dix). Throughout the book you will find shorthand web references of the form /e3/a-page-url/. Just prepend http://www.hcibook.com to find further information. To assist users of the second edition, a mapping between the structures of the old and new editions is available on the web at: http://www.hcibook.com/e3/contents/map2e/

Preface to the third edition **xxi**

STYLISTIC CONVENTION

As with all books, we have had to make some global decisions regarding style and terminology. Specifically, in a book in which the central characters are ‘the user’ and ‘the designer’, it is difficult to avoid the singular pronoun. We therefore use the pronoun ‘he’ when discussing the user and ‘she’ when referring to the designer. In other cases we use ‘she’ as a generic term. This should not be taken to imply anything about the composition of any actual population.

Similarly, we have adopted the convention of referring to the field of ‘Human– Computer Interaction’ and the notion of ‘human–computer interaction’. In many cases we will also use the abbreviation HCI.

ACKNOWLEDGEMENTS

In a book of this size, written by multiple authors, there will always be myriad people behind the scenes who have aided, supported and abetted our efforts. We would like to thank all those who provided information, pictures and software that have enhanced the quality of the final product. In particular, we are indebted to Wendy Mackay for the photograph of EVA; Wendy Hall and her colleagues at the University of Southampton for the screen shot of Microcosm; Saul Greenberg for the reactive keyboard; Alistair Edwards for Soundtrack; Christina Engelbart for the photographs of the early chord keyset and mouse; Geoff Ellis for the screen shot of Devina and himself using CuSeeMe; Steve Benford for images of the Internet Foyer; and Tony Renshaw who provided photographs of the eye tracking equipment. Thanks too to Simon Shum for information on design rationale, Robert Ward who gave us material on psycho-physiology, and Elizabeth Mynatt and Tom Rodden who worked with Gregory on material adapted in Chapter 20. Several of the boxed case studies are based on the work of multi-institution projects, and we are grateful

to all those from the project teams of CASCO, thePooch SMART-ITS, TOWER, AVATAR-Conference and TEAM-HOS for boxes and case studies based on their work; and also to the EQUATOR project from which we drew material for the boxes on cultural probes, ‘Ambient Wood’ and ‘City’. We would also like to thank all the reviewers and survey respondents whose feedback helped us to select our subject matter and improve our coverage; and our colleagues at our respective institutions and beyond who offered insight, encouragement and tolerance throughout the revi sion. We are indebted to all those who have contributed to the production process at Pearson Education and elsewhere, especially Keith Mansfield, Anita Atkinson, Lynette Miller, Sheila Chatten and Robert Chaundy.

Personal thanks must go to Fiona, Esther, Miriam, Rachel, Tina, Meghan, Aidan and Blaise, who have all endured ‘The Book’ well beyond the call of duty and over

**xxii** Preface to the third edition

many years, and Bruno and ‘the girls’ who continue to make their own inimitable contribution.

Finally we all owe huge thanks to Fiona for her continued deep personal support and for tireless proofreading, checking of figures, and keeping us all moving. We would never have got beyond the first edition without her.

The efforts of all of these have meant that the book is better than it would other wise have been. Where it could still be better, we take full responsibility.

PUBLISHER’S ACKNOWLEDGEMENTS

We are grateful to the following for permission to reproduce copyright material:

Figure p. 2, Figures 3.14, 3.15, 3.16 and 5.13 and Exercise 8.4 screen shots reprinted by permission from Apple Computer, Inc.; Figure 2.11 reprinted by permission of Keith Cheverst; Figure 3.13 from The WebBook and Web Forager: An information workspace for the world-wide web in *CHI Conference Proceedings*, © 1996 ACM, Inc., reprinted by permission (Card, S. K., Robertson, G. G. and York, W. 1996); Figures 3.9, 3.19, 5.5, Chapter 14, Design Focus: Looking real – Avatar Conference screen shots, Figures 21.3, 21.10, 21.11 screen shot frames reprinted by permission from Microsoft Corporation; Tables 6.2 and 6.3 adapted from Usability engineering: our experience and evolution in *Handbook for Human–Computer Interaction* edited by M. Helander, Copyright 1988, with permission from Elsevier (Whiteside, J., Bennett, J. and Hotzblatt, K. 1988); Figure 7.1 adapted from The alternate reality kit – an animated environment for creating interactive simulations in *Proceedings of Workshop on Visual Languages*, © 1986 IEEE, reprinted by permission of IEEE (Smith, R. B. 1986); Figure 7.2 from Guidelines for designing user interface software in *MITRE Corporation Report MTR-9420*, reprinted by permission of The MITRE Corporation (Smith, S. L. and Mosier, J. N. 1986); Figure 7.3 reprinted by permis sion of Jenifer Tidwell; Figures 8.6 and 8.9 from *Xview Programming Manual*, Volume 7 of *The X Window System*, reprinted by permission of O’Reilly and Associates, Inc. (Heller, D. 1990); Figure 9.8 screen shot reprinted by permission of Dr. R. D. Ward; Figure 10.2 after Earcons and icons: their structure and common design principles in *Human-Computer Interaction*, 4(1), published and reprinted by permission of Lawrence Erlbaum Associates, Inc. (Blattner, M., Sumikawa, D. and Greenberg, R. 1989); Figure 10.5 reprinted by permission of Alistair D. N. Edwards; Figure 10.7 reprinted by permission of Saul Greenberg; Figure 11.2 screen shot reprinted by permission of Macromedia, Inc.; Table 12.1 adapted from *The Psychology of Human Computer Interaction*, published and reprinted by permission of Lawrence Erlbaum Associates, Inc. (Card, S. K., Moran, T. P. and Newell, A. 1983); Table 12.2 after Table in A comparison of input devices in elemental pointing and dragging tasks in *Reaching through technology – CHI’91 Conference Proceedings, Human Factors in Computing Systems*, April, edited by S. P. Robertson, G. M. Olson and J. S. Olson, © 1991 ACM, Inc., reprinted by permission (Mackenzie,

**xxiv** Publisher’s acknowledgements

I. S., Sellen, A. and Buxton, W. 1991); Figure 14.1 from *Understanding Computers and Cognition: A New Foundation for Design*, published by Addison-Wesley, reprinted by permission of Pearson Education, Inc. (Winograd, T. and Flores, F. 1986); Figure 14.5 from *Theories of multi-party interaction. Technical report*, Social and Computer Sciences Research Group, University of Surrey and Queen Mary and Westfield Colleges, University of London, reprinted by permission of Nigel Gilbert (Hewitt, B., Gilbert, N., Jirotka, M. and Wilbur, S. 1990); Figure 14.6 from *Dialogue processes in computer-mediated communication: a study of letters in the com system. Technical report,* Linköping Studies in Arts and Sciences, reprinted by permission of Kerstin Severinson Eklundh (Eklundh, K. S. 1986); Chapter 14, Design Focus: Looking real – Avatar Conference, screen shots reprinted by permission of AVATAR-Conference project team; Figure 16.17 screen shot reprinted by permis sion of Harold Thimbleby; Figure 17.5 based on Verifying the behaviour of virtual world objects in *DSV-IS 2000 Interactive Systems: Design, Specification and Verification.* LNCS 1946, edited by P. Palanque and F. Paternò, published and reprinted by permission of Spinger-Verlag GmbH & Co. KG (Willans, J. S. and Harrison, M. D. 2001); Figure 18.4 icons reprinted by permission of Fabio Paternò; Chapter 19, p.675 CuSeeMe screen shot reprinted by permission of Geoff Ellis; Chapter 19, Design Focus: TOWER – workspace awareness, screen shots reprinted by permission of Wolfgang Prinz; Figure 20.1 reprinted by permission of Mitsubishi Electric Research Laboratories, Inc.; Figure 20.4 (right) reprinted by permission of Sony Computer Science Laboratories, Inc; Figure 20.9 from Cone trees. Animated 3d visualisation of hierarchical information in *Proceedings of the CH’91 Conference of Human Factors in Computing Systems*, © 1991 ACM, Inc., reprinted by permission (Robertson, G. G., Card, S. K., and Mackinlay, J. D. 1991); Figure 20.10 from Lifelines: visualising personal histories in *Proceedings of CH’96*, © 1996 ACM, Inc., reprinted by permission (Plaisant, C., Milash, B., Rose, A., Widoff, S. and Shneiderman, B. 1996); Figure 20.11 from Browsing anatomical image databases: a case study of the Visible Human in *CH’96 Conference Companion*, © 1996 ACM, Inc., reprinted by permission (North, C. and Korn, F. 1996); Figure 20.12 from Externalising abstract mathematical models in *Proceedings of CH’96*, © 1996 ACM, Inc., reprinted by permission (Tweedie, L., Spence, R., Dawkes, H. and Su, H. 1996); Figure 21.2 from The impact of Utility and Time on Distributed Information Retrieval in *People and Computers XII: Proceedings of HCI’97*, edited by H. Thimbleby, B. O’Conaill and P. Thomas, published and reprinted by permission of Spinger-Verlag GmbH & Co. KG (Johnson, C. W. 1997); Figure 21.4 screen shot reprinted by permission of the Departments of Electronics and Computer Science and History at the University of Southampton; Figure 21.6 Netscape browser window © 2002 Netscape Communications Corporation. Used with permission. Netscape has not authorized, sponsored, endorsed, or approved this publication and is not responsible for its content.

We are grateful to the following for permission to reproduce photographs:

Chapter 1, p. 50, Popperfoto.com; Chapter 2, p. 65, PCD Maltron Ltd; Figure 2.2 Electrolux; Figures 2.6 and 19.6 photos courtesy of Douglas Engelbart and Bootstrap Institute; Figure 2.8 (left) British Sky Broadcasting Limited; Figure 2.13 (bottom

Publisher’s acknowledgements **xxv**

right) Sony (UK) Ltd; Chapter 2, Design Focus: Feeling the Road, BMW AG; Chapter 2, Design Focus: Smart-Its – making using sensors easy, Hans Gellersen; Figures 4.1 (right) and 20.2 (left) Palo Alto Research Center; Figure 4.2 and 20.3 (left) François Guimbretière; Figure 4.3 (bottom left) Franklin Electronic Publishers; Figure 5.2 (top plate and middle plate) Kingston Museum and Heritage Service, (bottom plate) V&A Images, The Victoria and Albert Museum, London; Chapter 5, Design Focus: Cultural probes, William W. Gaver, Anthony Boucher, Sarah Pennington and Brendan Walker, Equator IRC, Royal College of Art; Chapter 6, p. 245, from The 1984 Olympic Message System: a text of behavioural principle of system design in *Communications of the ACM*, 30(9), © 1987 ACM, Inc., reprinted by permission (Gould, J. D., Boies, S. J., Levy, S., Richards, J. T. and Schoonard, J. 1987); Figures 9.5 and 9.6 J. A. Renshaw; Figure 9.7 Dr. R. D. Ward; Figure 10.3 SensAble Technologies; Chapter 13, Design Focus: Tomorrow’s hospital – using

participatory design, Professor J. Artur Vale Serrano; Chapter 18, p. 650, Michael Beigl; Chapter 19, p. 678, Steve Benford, The Mixed Reality Laboratory, University of Nottingham; Chapter 19, Design Focus: SMS in action, Mark Rouncefield; Figure 20.2 (right) Ken Hinckley; Figure 20.3 (right) MIT Media Lab; Figure 20.4 (left) from Interacting with paper on the digital desk in *Communications of the ACM*, 36(7), © 1993 ACM, Inc., reprinted by permission (Wellner, P. 1993); Chapter 20,

p. 726, Peter Phillips; Chapter 20, Design Focus: Ambient wood – augmenting the physical, Yvonne Rogers; Chapter 20, Design Focus: Shared experience, Matthew Chalmers.

We are grateful to the following for permission to reproduce text extracts:

Pearson Education, Inc. Publishing as Pearson Addison Wesley for an extract adapted from *Designing the User Interface: Strategies for Effective Human–Computer Interaction 3/e* by B. Shneiderman © 1998, Pearson Education, Inc; Perseus Books Group for an extract adapted from *The Design of Everyday Things* by D. Norman, 1998; and Wiley Publishing, Inc. for extracts adapted from ‘Heuristic Evaluation’ by Jakob Nielson and Robert L. Mack published in *Usability Inspection Methods* © 1994 Wiley Publishing, Inc.; IEEE for permission to base chapter 20 on ‘The human ex perience’ by Gregory Abowd, Elizabeth Mynatt and Tom Rodden which appeared in *IEEE Pervasive Computing Magazine*, Special Inaugural Issue on Reaching for Weiser’s Vision, Vol. 1, Issue 1, pp. 48–58, Jan–March 2002. © 2002 IEEE.

In some instances we have been unable to trace the owners of copyright material, and we would appreciate any information that would enable us to do so.

INTRODUCTION

WHY HUMAN–COMPUTER INTERACTION?

In the first edition of this book we wrote the following:

This is the authors’ second attempt at writing this introduction. Our first attempt fell victim to a design quirk coupled with an innocent, though weary and less than attentive, user. The word-processing package we originally used to write this intro duction is menu based. Menu items are grouped to reflect their function. The ‘save’ and ‘delete’ options, both of which are correctly classified as file-level operations, are consequently adjacent items in the menu. With a cursor controlled by a trackball it is all too easy for the hand to slip, inadvertently selecting delete instead of save. Of course, the delete option, being well thought out, pops up a confirmation box allow ing the user to cancel a mistaken command. Unfortunately, the save option produces a very similar confirmation box – it was only as we hit the ‘Confirm’ button that we noticed the word ‘delete’ at the top...

Happily this word processor no longer has a delete option in its menu, but unfortu nately, similar problems to this are still an all too common occurrence. Errors such as these, resulting from poor design choices, happen every day. Perhaps they are not catastrophic: after all nobody’s life is endangered nor is there environmental damage (unless the designer happens to be nearby or you break something in frustration!). However, when you lose several hours’ work with no written notes or backup and

a publisher’s deadline already a week past, ‘catastrophe’ is certainly the word that springs to mind.

Why is it then that when computers are marketed as ‘user friendly’ and ‘easy to use’, simple mistakes like this can still occur? Did the designer of the word processor actually try to use it with the trackball, or was it just that she was so expert with the system that the mistake never arose? We hazard a guess that no one tried to use it when tired and under pressure. But these criticisms are not levied only on the design ers of traditional computer software. More and more, our everyday lives involve pro grammed devices that do not sit on our desk, and these devices are just as unusable. Exactly how many VCR designers understand the universal difficulty people have trying to set their machines to record a television program? Do car radio designers

**2** Introduction

actually think it is safe to use so many knobs and displays that the driver has to divert attention away from the road completely in order to tune the radio or adjust the volume?

Computers and related devices have to be designed with an understanding that people with specific tasks in mind will want to use them in a way that is seamless with respect to their everyday work. To do this, those who design these systems need to know how to think in terms of the eventual users’ tasks and how to translate that knowledge into an executable system. But there is a problem with trying to teach the notion of designing computers for people. All designers *are* people and, most prob ably, they are users as well. Isn’t it therefore intuitive to design for the user? Why does it need to be taught when we all know what a good interface looks like? As a result, the study of human–computer interaction (HCI) tends to come late in the designer’s training, if at all. The scenario with which we started shows that this is a mistaken view; it is not at all intuitive or easy to design consistent, robust systems

DESIGN FOCUS 

Things don’t change

It would be nice to think that problems like those described at the start of the Introduction would never happen now. Think again! Look at the MacOS X ‘dock’ below. It is a fast launch point for applica tions; folders and files can be dragged there for instant access; and also, at the right-hand side, there sits the trash can. Imagine what happens as you try to drag a file into one of the folders. If your finger accidentally slips whilst the icon is over the trash can – oops!

Happily this is not quite as easy in reality as it looks in the screen shot, since the icons in the dock con stantly move around as you try to drag a file into it. This is to make room for the file in case you want to place it in the dock. However, it means you have to concentrate very hard when dragging a file over the dock. We assume this is not a deliberate feature, but it does have the beneficial side effect that users are less likely to throw away a file by accident – whew!

In fact it is quite fun to watch a new user trying to throw away a file. The trash can keeps moving as if it didn’t want the file in it. Experienced users evolve coping strategies. One user always drags files into the trash from the right-hand side as then the icons in the dock don’t move around. So two lessons:

n designs don’t always get better

n but at least users are clever.

Screen shot reprinted by permission from Apple Computer, Inc.

What is HCI? **3**

that will cope with all manner of user carelessness. The interface is not something that can be plugged in at the last minute; its design should be developed integrally with the rest of the system. It should not just present a ‘pretty face’, but should sup port the tasks that people actually want to do, and forgive the careless mistakes. We therefore need to consider how HCI fits into the design process.

Designing usable systems is not simply a matter of altruism towards the eventual user, or even marketing; it is increasingly a matter of law. National health and safety standards constrain employers to provide their workforce with usable computer sys tems: not just safe but *usable*. For example, EC Directive 90/270/EEC, which has been incorporated into member countries’ legislation, requires employers to ensure the following when designing, selecting, commissioning or modifying software:

n that it is suitable for the task

n that it is easy to use and, where appropriate, adaptable to the user’s knowledge and experience

n that it provides feedback on performance

n that it displays information in a format and at a pace that is adapted to the user n that it conforms to the ‘principles of software ergonomics’.

Designers and employers can no longer afford to ignore the user.

WHAT IS HCI?

The term *human–computer interaction* has only been in widespread use since the early 1980s, but has its roots in more established disciplines. Systematic study of human performance began in earnest at the beginning of the last century in factories, with an emphasis on manual tasks. The Second World War provided the impetus for studying the interaction between humans and machines, as each side strove to pro duce more effective weapons systems. This led to a wave of interest in the area among researchers, and the formation of the Ergonomics Research Society in 1949. Tradi tionally, ergonomists have been concerned primarily with the physical characteristics of machines and systems, and how these affect user performance. Human Factors incorporates these issues, and more cognitive issues as well. The terms are often used interchangeably, with Ergonomics being the preferred term in the United Kingdom and Human Factors in the English-speaking parts of North America. Both of these disciplines are concerned with user performance in the context of any system, whether computer, mechanical or manual. As computer use became more widespread, an increasing number of researchers specialized in studying the interaction between people and computers, concerning themselves with the physical, psychological and theoretical aspects of this process. This research originally went under the name *man– machine interaction*, but this became *human–computer interaction* in recognition of the particular interest in computers and the composition of the user population!

Another strand of research that has influenced the development of HCI is infor mation science and technology. Again the former is an old discipline, pre-dating the introduction of technology, and is concerned with the management and manipulation

**4** Introduction

of information within an organization. The introduction of technology has had a profound effect on the way that information can be stored, accessed and utilized and, consequently, a significant effect on the organization and work environment. Systems analysis has traditionally concerned itself with the influence of technology in the workplace, and fitting the technology to the requirements and constraints of the job. These issues are also the concern of HCI.

HCI draws on many disciplines, as we shall see, but it is in computer science and systems design that it must be accepted as a central concern. For all the other discip lines it can be a specialism, albeit one that provides crucial input; for systems design it is an essential part of the design process. From this perspective, HCI involves the design, implementation and evaluation of interactive systems in the context of the user’s task and work.

However, when we talk about human–computer interaction, we do not necessarily envisage a single user with a desktop computer. By *user* we may mean an individual user, a group of users working together, or a sequence of users in an organization, each dealing with some part of the task or process. The user is whoever is trying to get the job done using the technology. By *computer* we mean any technology ranging from the general desktop computer to a large-scale computer system, a process control system or an embedded system. The system may include non-computerized parts, including other people. By *interaction* we mean any communication between a user and computer, be it direct or indirect. Direct interaction involves a dialog

with feedback and control throughout performance of the task. Indirect interaction may involve batch processing or intelligent sensors controlling the environment. The important thing is that the user is interacting with the computer in order to accomplish something.

WHO IS INVOLVED IN HCI?

HCI is undoubtedly a multi-disciplinary subject. The ideal designer of an interactive system would have expertise in a range of topics: psychology and cognitive science to give her knowledge of the user’s perceptual, cognitive and problem-solving skills; ergonomics for the user’s physical capabilities; sociology to help her under stand the wider context of the interaction; computer science and engineering to be able to build the necessary technology; business to be able to market it; graphic design to produce an effective interface presentation; technical writing to produce the manuals, and so it goes on. There is obviously too much expertise here to be held by one person (or indeed four!), perhaps even too much for the average design team. Indeed, although HCI is recognized as an interdisciplinary subject, in practice peo ple tend to take a strong stance on one side or another. However, it is not possible to design effective interactive systems from one discipline in isolation. Input is needed from all sides. For example, a beautifully designed graphic display may be unusable if it ignores dialog constraints or the psychological limitations of the user.

Theory and HCI **5**

In this book we want to encourage the multi-disciplinary view of HCI but we too have our ‘stance’, as computer scientists. We are interested in answering a particular question. How do principles and methods from each of these contributing dis ciplines in HCI help us to design better systems? In this we must be pragmatists rather than theorists: we want to know how to apply the theory to the problem

rather than just acquire a deep understanding of the theory. Our goal, then, is to be multi-disciplinary but practical. We concentrate particularly on computer science, psychology and cognitive science as core subjects, and on their application to design; other disciplines are consulted to provide input where relevant.

THEORY AND HCI

Unfortunately for us, there is no general and unified theory of HCI that we can present. Indeed, it may be impossible ever to derive one; it is certainly out of our reach today. However, there is an underlying principle that forms the basis of our own views on HCI, and it is captured in our claim that people use computers to accomplish work. This outlines the three major issues of concern: the people, the computers and the tasks that are performed. The system must support the user’s task, which gives us a fourth focus, usability: if the system forces the user to adopt an unacceptable mode of work then it is not usable.

There are, however, those who would dismiss our concentration on the task, saying that we do not even know enough about a theory of human tasks to support them in design. There is a good argument here (to which we return in Chapter 15). However, we can live with this confusion about what real tasks are because our understanding of tasks at the moment is sufficient to give us direction in design. The user’s current tasks are studied and then supported by computers, which can in turn affect the nature of the original task and cause it to evolve. To illustrate, word processing has made it easy to manipulate paragraphs and reorder documents, allowing writers a completely new freedom that has affected writing styles. No longer is it vital to plan and construct text in an ordered fashion, since free-flowing prose can easily be restructured at a later date. This evolution of task in turn affects the design of the ideal system. However, we see this evolution as providing a motivating force behind the system development cycle, rather than a refutation of the whole idea of supportive design.

This word ‘task’ or the focus on accomplishing ‘work’ is also problematic when we think of areas such as domestic appliances, consumer electronics and e-commerce. There are three ‘use’ words that must all be true for a product to be successful; it must be:

**useful** – accomplish what is required: play music, cook dinner, format a document; **usable** – do it easily and naturally, without danger of error, etc.; **used** – make people want to use it, be attractive, engaging, fun, etc.

**6** Introduction

The last of these has not been a major factor until recently in HCI, but issues of motivation, enjoyment and experience are increasingly important. We are certainly even further from having a unified theory of experience than of task.

The question of whether HCI, or more importantly the design of interactive sys tems and the user interface in particular, is a science or a craft discipline is an inter esting one. Does it involve artistic skill and fortuitous insight or reasoned methodical science? Here we can draw an analogy with architecture. The most impressive struc tures, the most beautiful buildings, the innovative and imaginative creations that provide aesthetic pleasure, all require inventive inspiration in design and a sense of artistry, and in this sense the discipline is a craft. However, these structures also have to be able to stand up to fulfill their purpose successfully, and to be able to do this the architect has to use science. So it is for HCI: beautiful and/or novel interfaces are artistically pleasing *and* capable of fulfilling the tasks required – a marriage of art and science into a successful whole. We want to reuse lessons learned from the past about how to achieve good results and avoid bad ones. For this we require both craft and science. Innovative ideas lead to more usable systems, but in order to maximize the potential benefit from the ideas, we need to understand not only that they work, but how and why they work. This scientific rationalization allows us to reuse related con cepts in similar situations, in much the same way that architects can produce a bridge and know that it will stand, since it is based upon tried and tested principles.

The craft–science tension becomes even more difficult when we consider novel systems. Their increasing complexity means that our personal ideas of good and bad are no longer enough; for a complex system to be well designed we need to rely on something more than simply our intuition. Designers may be able to think about how one user would want to act, but how about groups? And what about new media? Our ideas of how best to share workloads or present video information are open to debate and question even in non-computing situations, and the incorporation of one version of good design into a computer system is quite likely to be unlike anyone else’s version. Different people work in different ways, whilst different media color the nature of the interaction; both can dramatically change the very nature of the original task. In order to assist designers, it is unrealistic to assume that they can rely on artistic skill and perfect insight to develop usable systems. Instead we have to pro vide them with an understanding of the concepts involved, a scientific view of the reasons why certain things are successful whilst others are not, and then allow their creative nature to feed off this information: creative flow, underpinned with science; or maybe scientific method, accelerated by artistic insight. The truth is that HCI is required to be both a craft and a science in order to be successful.

HCI IN THE CURRICULUM

If HCI involves both craft and science then it must, in part at least, be taught. Imagination and skill may be qualities innate in the designer or developed through experience, but the underlying theory must be learned. In the past, when computers

HCI in the curriculum **7**

were used primarily by expert specialists, concentration on the interface was a lux ury that was often relinquished. Now designers cannot afford to ignore the interface in favour of the functionality of their systems: the two are too closely intertwined. If the interface is poor, the functionality is obscured; if it is well designed, it will allow the system’s functionality to support the user’s task.

Increasingly, therefore, computer science educators cannot afford to ignore HCI. We would go as far as to claim that HCI should be integrated into every computer science or software engineering course, either as a recurring feature of other modules or, preferably, as a module itself. It should not be viewed as an ‘optional extra’ (although, of course, more advanced HCI options can complement a basic core course). This view is shared by the ACM SIGCHI curriculum development group, who propose a curriculum for such a core course [9]. The topics included in this book, although developed without reference to this curriculum, cover the main emphases of it, and include enough detail and coverage to support specialized options as well.

In courses other than computer science, HCI may well be an option specializing in a particular area, such as cognitive modeling or task analysis. Selected use of the relevant chapters of this book can also support such a course.

HCI must be taken seriously by designers and educators if the requirement for additional complexity in the system is to be matched by increased clarity and usabil ity in the interface. In this book we demonstrate how this can be done in practice.

DESIGN FOCUS 

Quick fixes

You should expect to spend both time and money on interface design, just as you would with other parts of a system. So in one sense there are no quick fixes. However, a few simple steps can make a dramatic improvement.

**Think ‘user’**

Probably 90% of the value of any interface design technique is that it forces the designer to remember that someone (and in particular someone else) will use the system under construction.

**Try it out**

Of course, many designers will build a system that they find easy and pleasant to use, and they find it incomprehensible that anyone else could have trouble with it. Simply sitting someone down with an early version of an interface (without the designer prompting them at each step!) is enormously valuable. Professional usability laboratories will have video equipment, one-way mirrors and other

sophisticated monitors, but a notebook and pencil and a home-video camera will suffice (more about evaluation in Chapter 9).

**Involve the users**

Where possible, the eventual users should be involved in the design process. They have vital know ledge and will soon find flaws. A mechanical syringe was once being developed and a prototype was demonstrated to hospital staff. Happily they quickly noticed the potentially fatal flaw in its interface.

**8** Introduction

**Figure 0.1** Automatic syringe: setting the dose to 1372. The effect of one key slip before and after user involvement

The doses were entered via a numeric keypad: an accidental keypress and the dose could be out by a factor of 10! The production version had individual increment/decrement buttons for each digit (more about participatory design in Chapter 13).

**Iterate**

People are complicated, so you won’t get it right first time. Programming an interface can be a very difficult and time-consuming business. So, the result becomes precious and the builder will want to defend it and minimize changes. Making early prototypes less precious and easier to throw away is crucial. Happily there are now many interface builder tools that aid this process. For example, mock ups can be quickly constructed using HyperCard on the Apple Macintosh or Visual Basic on the PC. For visual and layout decisions, paper designs and simple models can be used (more about iterative design in Chapter 5).

PART

FOUNDATIONS 1

In this part we introduce the fundamental components of

an interactive system: the human user, the computer system

itself and the nature of the interactive process. We then

present a view of the history of interactive systems by look

ing at key interaction paradigms that have been significant.

Chapter 1 discusses the psychological and physiological

attributes of the user, providing us with a basic overview of

the capabilities and limitations that affect our ability to use

computer systems. It is only when we have an understand

ing of the user at this level that we can understand what

makes for successful designs. Chapter 2 considers the

computer in a similar way. Input and output devices are

described and explained and the effect that their individual

characteristics have on the interaction highlighted. The

computational power and memory of the computer is

another important component in determining what can be

achieved in the interaction, whilst due attention is also paid

to paper output since this forms one of the major uses

of computers and users’ tasks today. Having approached

interaction from both the human and the computer side,

we then turn our attention to the dialog between them

in Chapter 3, where we look at models of interaction. In

Chapter 4 we take a historical perspective on the evolution

of interactive systems and how they have increased the

usability of computers in general.

THE HUMAN 1

OVERVIEW

n Humans are limited in their capacity to process

information. This has important implications for design.

n Information is received and responses given via a

number of input and output channels:

– visual channel

– auditory channel

– haptic channel

– movement.

n Information is stored in memory:

– sensory memory

– short-term (working) memory

– long-term memory.

n Information is processed and applied:

– reasoning

– problem solving

– skill acquisition

– error.

n Emotion influences human capabilities.

n Users share common capabilities but are individuals

with differences, which should not be ignored.

**12** Chapter 1 n The human

**1.1**

INTRODUCTION

This chapter is the first of four in which we introduce some of the ‘foundations’ of HCI. We start with the human, the central character in any discussion of interactive systems. The human, the *user*, is, after all, the one whom computer systems are de signed to assist. The requirements of the user should therefore be our first priority.

In this chapter we will look at areas of human psychology coming under the general banner of *cognitive psychology*. This may seem a far cry from designing and building interactive computer systems, but it is not. In order to design something for some one, we need to understand their capabilities and limitations. We need to know if there are things that they will find difficult or, even, impossible. It will also help us to know what people find easy and how we can help them by encouraging these things. We will look at aspects of cognitive psychology which have a bearing on the use of com puter systems: how humans perceive the world around them, how they store and process information and solve problems, and how they physically manipulate objects.

We have already said that we will restrict our study to those things that are relev ant to HCI. One way to structure this discussion is to think of the user in a way that highlights these aspects. In other words, to think of a simplified *model* of what is actually going on. Many models have been proposed and it useful to consider one of the most influential in passing, to understand the context of the discussion that is to follow. In 1983, Card, Moran and Newell [56] described the *Model Human Processor*, which is a simplified view of the human processing involved in interacting with computer systems. The model comprises three subsystems: the perceptual system, handling sensory stimulus from the outside world, the motor system, which controls actions, and the cognitive system, which provides the processing needed to connect the two. Each of these subsystems has its own processor and memory, although

obviously the complexity of these varies depending on the complexity of the tasks the subsystem has to perform. The model also includes a number of *principles of operation* which dictate the behavior of the systems under certain conditions.

We will use the analogy of the user as an information processing system, but in our model make the analogy closer to that of a conventional computer system. Information comes in, is stored and processed, and information is passed out. We will therefore discuss three components of this system: input–output, memory and processing. In the human, we are dealing with an intelligent information-processing system, and processing therefore includes problem solving, learning, and, con sequently, making mistakes. This model is obviously a simplification of the real

situation, since memory and processing are required at all levels, as we have seen in the Model Human Processor. However, it is convenient as a way of grasping how information is handled by the human system. The human, unlike the computer, is also influenced by external factors such as the social and organizational environ ment, and we need to be aware of these influences as well. We will ignore such

factors for now and concentrate on the human’s information processing capabilities only. We will return to social and organizational influences in Chapter 3 and, in more detail, in Chapter 13.

1.2 Input–output channels **13**

In this chapter, we will first look at the human’s input–output channels, the senses and responders or effectors. This will involve some low-level processing. Secondly, we will consider human memory and how it works. We will then think about how humans perform complex problem solving, how they learn and acquire skills, and why they make mistakes. Finally, we will discuss how these things can help us in the design of computer systems.

**1.2**

INPUT–OUTPUT CHANNELS

A person’s interaction with the outside world occurs through information being received and sent: input and output. In an interaction with a computer the user receives information that is output by the computer, and responds by providing input to the computer – the user’s output becomes the computer’s input and vice versa. Consequently the use of the terms input and output may lead to confusion so we shall blur the distinction somewhat and concentrate on the channels involved. This blurring is appropriate since, although a particular channel may have a primary role as input or output in the interaction, it is more than likely that it is also used in the other role. For example, sight may be used primarily in receiving information from the computer, but it can also be used to provide information to the computer, for example by fixating on a particular screen point when using an eyegaze system.

Input in the human occurs mainly through the senses and output through the motor control of the effectors. There are five major senses: sight, hearing, touch, taste and smell. Of these, the first three are the most important to HCI. Taste and smell do not currently play a significant role in HCI, and it is not clear whether they could be exploited at all in general computer systems, although they could have a role to play in more specialized systems (smells to give warning of malfunction, for example) or in augmented reality systems. However, vision, hearing and touch are central.

Similarly there are a number of effectors, including the limbs, fingers, eyes, head and vocal system. In the interaction with the computer, the fingers play the primary role, through typing or mouse control, with some use of voice, and eye, head and body position.

Imagine using a personal computer (PC) with a mouse and a keyboard. The appli cation you are using has a graphical interface, with menus, icons and windows. In your interaction with this system you receive information primarily by sight, from what appears on the screen. However, you may also receive information by ear: for example, the computer may ‘beep’ at you if you make a mistake or to draw attention to something, or there may be a voice commentary in a multimedia presentation. Touch plays a part too in that you will feel the keys moving (also hearing the ‘click’) or the orientation of the mouse, which provides vital feedback about what you have done. You yourself send information to the computer using your hands, either by hitting keys or moving the mouse. Sight and hearing do not play a direct role

in sending information in this example, although they may be used to receive

**14** Chapter 1 n The human

information from a third source (for example, a book, or the words of another per son) which is then transmitted to the computer.

In this section we will look at the main elements of such an interaction, first con sidering the role and limitations of the three primary senses and going on to consider motor control.

1.2.1 Vision

Human vision is a highly complex activity with a range of physical and perceptual limitations, yet it is the primary source of information for the average person. We can roughly divide visual perception into two stages: the physical reception of the stimulus from the outside world, and the processing and interpretation of that stimulus. On the one hand the physical properties of the eye and the visual system mean that there are certain things that cannot be seen by the human; on the other the interpretative capabilities of visual processing allow images to be constructed from incomplete information. We need to understand both stages as both influence what can and cannot be perceived visually by a human being, which in turn directly affects the way that we design computer systems. We will begin by looking at the

eye as a physical receptor, and then go on to consider the processing involved in basic vision.

*The human eye*

Vision begins with light. The eye is a mechanism for receiving light and transform ing it into electrical energy. Light is reflected from objects in the world and their image is focussed upside down on the back of the eye. The receptors in the eye transform it into electrical signals which are passed to the brain.

The eye has a number of important components (see Figure 1.1) which we will look at in more detail. The *cornea* and *lens* at the front of the eye focus the light into a sharp image on the back of the eye, the *retina*. The retina is light sensitive and con tains two types of *photoreceptor*: *rods* and *cones*.

Rods are highly sensitive to light and therefore allow us to see under a low level of illumination. However, they are unable to resolve fine detail and are subject to light saturation. This is the reason for the temporary blindness we get when moving from a darkened room into sunlight: the rods have been active and are saturated by the sudden light. The cones do not operate either as they are suppressed by the rods. We are therefore temporarily unable to see at all. There are approximately 120 million rods per eye which are mainly situated towards the edges of the retina. Rods there fore dominate peripheral vision.

Cones are the second type of receptor in the eye. They are less sensitive to light than the rods and can therefore tolerate more light. There are three types of cone, each sensitive to a different wavelength of light. This allows color vision. The eye has approximately 6 million cones, mainly concentrated on the *fovea*, a small area of the retina on which images are fixated.

**Figure 1.1** The human eye

1.2 Input–output channels **15**

Although the retina is mainly covered with photoreceptors there is one *blind spot* where the optic nerve enters the eye. The blind spot has no rods or cones, yet our visual system compensates for this so that in normal circumstances we are unaware of it.

The retina also has specialized nerve cells called *ganglion cells*. There are two types: X-cells, which are concentrated in the fovea and are responsible for the early detec tion of pattern; and Y-cells which are more widely distributed in the retina and are responsible for the early detection of movement. The distribution of these cells means that, while we may not be able to detect changes in pattern in peripheral vision, we can perceive movement.

*Visual perception*

Understanding the basic construction of the eye goes some way to explaining the physical mechanisms of vision but visual perception is more than this. The informa tion received by the visual apparatus must be filtered and passed to processing ele ments which allow us to recognize coherent scenes, disambiguate relative distances and differentiate color. We will consider some of the capabilities and limitations of visual processing later, but first we will look a little more closely at how we perceive size and depth, brightness and color, each of which is crucial to the design of effective visual interfaces.

**16** Chapter 1 n The human

DESIGN FOCUS 

Getting noticed

The extensive knowledge about the human visual system can be brought to bear in practical design. For example, our ability to read or distinguish falls off inversely as the distance from our point of focus increases. This is due to the fact that the cones are packed more densely towards the center of our visual field. You can see this in the following image. Fixate on the dot in the center. The letters on the left should all be equally readable, those on the right all equally harder.

This loss of discrimination sets limits on the amount that can be seen or read without moving one’s eyes. A user concentrating on the middle of the screen cannot be expected to read help text on the bottom line.

However, although our ability to discriminate static text diminishes, the rods, which are concentrated more in the outer parts of our visual field, are very sensitive to changes; hence we see movement well at the edge of our vision. So if you want a user to see an error message at the bottom of the screen it had better be flashing! On the other hand clever moving icons, however impressive they are, will be distracting even when the user is not looking directly at them.

***Perceiving size and depth*** Imagine you are standing on a hilltop. Beside you on the summit you can see rocks, sheep and a small tree. On the hillside is a farmhouse with outbuildings and farm vehicles. Someone is on the track, walking toward the summit. Below in the valley is a small market town.

Even in describing such a scene the notions of size and distance predominate. Our visual system is easily able to interpret the images which it receives to take account of these things. We can identify similar objects regardless of the fact that they appear to us to be of vastly different sizes. In fact, we can use this information to judge

distances.

So how does the eye perceive size, depth and relative distances? To understand this we must consider how the image appears on the retina. As we noted in the previous section, reflected light from the object forms an upside-down image on the retina. The size of that image is specified as a *visual angle*. Figure 1.2 illustrates how the visual angle is calculated.

If we were to draw a line from the top of the object to a central point on the front of the eye and a second line from the bottom of the object to the same point, the visual angle of the object is the angle between these two lines. Visual angle is affected by both the size of the object and its distance from the eye. Therefore if two objects are at the same distance, the larger one will have the larger visual angle. Similarly, if two objects of the same size are placed at different distances from the eye, the

**Figure 1.2** Visual angle

1.2 Input–output channels **17**

furthest one will have the smaller visual angle. The visual angle indicates how much of the field of view is taken by the object. The visual angle measurement is given in either degrees or *minutes of arc*, where 1 degree is equivalent to 60 minutes of arc, and 1 minute of arc to 60 seconds of arc.

So how does an object’s visual angle affect our perception of its size? First, if the visual angle of an object is too small we will be unable to perceive it at all. *Visual acuity* is the ability of a person to perceive fine detail. A number of measurements have been established to test visual acuity, most of which are included in standard eye tests. For example, a person with normal vision can detect a single line if it has a visual angle of 0.5 seconds of arc. Spaces between lines can be detected at 30 seconds to 1 minute of visual arc. These represent the limits of human visual acuity.

Assuming that we can perceive the object, does its visual angle affect our per ception of its size? Given that the visual angle of an object is reduced as it gets further away, we might expect that we would perceive the object as smaller. In fact, our perception of an object’s size remains constant even if its visual angle changes. So a person’s height is perceived as constant even if they move further from you. This is the *law of size constancy*, and it indicates that our perception of size relies on factors other than the visual angle.

One of these factors is our perception of depth. If we return to the hilltop scene there are a number of *cues* which we can use to determine the relative positions and distances of the objects which we see. If objects overlap, the object which is partially covered is perceived to be in the background, and therefore further away. Similarly, the size and height of the object in our field of view provides a cue to its distance.

**18** Chapter 1 n The human

A third cue is familiarity: if we expect an object to be of a certain size then we can judge its distance accordingly. This has been exploited for humour in advertising: one advertisement for beer shows a man walking away from a bottle in the fore ground. As he walks, he bumps into the bottle, which is in fact a giant one in the background!

***Perceiving brightness*** A second aspect of visual perception is the perception of *brightness*. Brightness is in fact a subjective reaction to levels of light. It is affected by *luminance* which is the amount of light emitted by an object. The luminance of an object is dependent on the amount of light falling on the object’s surface and its reflective properties. Luminance is a physical characteristic and can be measured using a *photometer*. *Contrast* is related to luminance: it is a function of the luminance of an object and the luminance of its background.

Although brightness is a subjective response, it can be described in terms of the amount of luminance that gives a *just noticeable difference* in brightness. However, the visual system itself also compensates for changes in brightness. In dim lighting, the rods predominate vision. Since there are fewer rods on the fovea, objects in low lighting can be seen less easily when fixated upon, and are more visible in peripheral vision. In normal lighting, the cones take over.

Visual acuity increases with increased luminance. This may be an argument for using high display luminance. However, as luminance increases, *flicker* also increases. The eye will perceive a light switched on and off rapidly as constantly on. But if the speed of switching is less than 50 Hz then the light is perceived to flicker. In high luminance flicker can be perceived at over 50 Hz. Flicker is also more noticeable in peripheral vision. This means that the larger the display (and consequently the more peripheral vision that it occupies), the more it will appear to flicker.

***Perceiving color*** A third factor that we need to consider is perception of color. Color is usually regarded as being made up of three components: *hue*, *intensity* and *saturation*. Hue is determined by the spectral wavelength of the light. Blues have short wavelengths, greens medium and reds long. Approximately 150 different hues can be discriminated by the average person. Intensity is the brightness of the color, and

saturation is the amount of whiteness in the color. By varying these two, we can perceive in the region of 7 million different colors. However, the number of colors that can be identified by an individual without training is far fewer (in the region of 10).

The eye perceives color because the cones are sensitive to light of different wave lengths. There are three different types of cone, each sensitive to a different color (blue, green and red). Color vision is best in the fovea, and worst at the periphery where rods predominate. It should also be noted that only 3–4% of the fovea is occupied by cones which are sensitive to blue light, making blue acuity lower.

Finally, we should remember that around 8% of males and 1% of females suffer from color blindness, most commonly being unable to discriminate between red and green.

1.2 Input–output channels **19**

*The capabilities and limitations of visual processing*

In considering the way in which we perceive images we have already encountered some of the capabilities and limitations of the human visual processing system. However, we have concentrated largely on low-level perception. Visual processing involves the transformation and interpretation of a complete image, from the light that is thrown onto the retina. As we have already noted, our expectations affect the way an image is perceived. For example, if we know that an object is a particular size, we will perceive it as that size no matter how far it is from us.

Visual processing compensates for the movement of the image on the retina which occurs as we move around and as the object which we see moves. Although the retinal image is moving, the image that we perceive is stable. Similarly, color and brightness of objects are perceived as constant, in spite of changes in luminance.

This ability to interpret and exploit our expectations can be used to resolve ambi guity. For example, consider the image shown in Figure 1.3. What do you perceive? Now consider Figure 1.4 and Figure 1.5. The context in which the object appears

**Figure 1.3** An ambiguous shape?

**Figure 1.4** ABC

**20** Chapter 1 n The human

**Figure 1.5** 12 13 14

**Figure 1.6** The Muller–Lyer illusion – which line is longer?

allows our expectations to clearly disambiguate the interpretation of the object, as either a B or a 13.

However, it can also create optical illusions. For example, consider Figure 1.6. Which line is longer? Most people when presented with this will say that the top line is longer than the bottom. In fact, the two lines are the same length. This may be due to a false application of the law of size constancy: the top line appears like a con cave edge, the bottom like a convex edge. The former therefore seems further away than the latter and is therefore scaled to appear larger. A similar illusion is the Ponzo illusion (Figure 1.7). Here the top line appears longer, owing to the distance effect, although both lines are the same length. These illusions demonstrate that our per ception of size is not completely reliable.

Another illusion created by our expectations compensating an image is the proof reading illusion. Read the text in Figure 1.8 quickly. What does it say? Most people reading this rapidly will read it correctly, although closer inspection shows that the word ‘the’ is repeated in the second and third line.

These are just a few examples of how the visual system compensates, and some times overcompensates, to allow us to perceive the world around us.

1.2 Input–output channels **21**

**Figure 1.7** The Ponzo illusion – are these the same size?

**Figure 1.8** Is this text correct?

**22** Chapter 1 n The human

DESIGN FOCUS 

Where’s the middle?

Optical illusions highlight the differences between the way things are and the way we perceive them – and in interface design we need to be aware that we will not always perceive things exactly as they are. The way that objects are composed together will affect the way we perceive them, and we do not per ceive geometric shapes exactly as they are drawn. For example, we tend to magnify horizontal lines and reduce vertical. So a square needs to be slightly increased in height to appear square and lines will appear thicker if horizontal rather than vertical.

Optical illusions also affect page symmetry. We tend to see the center of a page as being a little above the actual center – so if a page is arranged symmetrically around the actual center, we will see it as too low down. In graphic design this is known as the *optical center* – and bottom page margins tend to be increased by 50% to compensate.

*Reading*

So far we have concentrated on the perception of images in general. However, the perception and processing of text is a special case that is important to interface design, which invariably requires some textual display. We will therefore end this section by looking at *reading*. There are several stages in the reading process. First, the visual pattern of the word on the page is perceived. It is then decoded with reference to an internal representation of language. The final stages of lan guage processing include syntactic and semantic analysis and operate on phrases or sentences.

We are most concerned with the first two stages of this process and how they influence interface design. During reading, the eye makes jerky movements called *saccades* followed by fixations. Perception occurs during the fixation periods, which account for approximately 94% of the time elapsed. The eye moves backwards over the text as well as forwards, in what are known as *regressions*. If the text is complex there will be more regressions.

Adults read approximately 250 words a minute. It is unlikely that words are scanned serially, character by character, since experiments have shown that words can be recognized as quickly as single characters. Instead, familiar words are recognized using word shape. This means that removing the word shape clues (for example, by capitalizing words) is detrimental to reading speed and accuracy.

The speed at which text can be read is a measure of its legibility. Experiments have shown that standard font sizes of 9 to 12 points are equally legible, given pro portional spacing between lines [346]. Similarly line lengths of between 2.3 and 5.2 inches (58 and 132 mm) are equally legible. However, there is evidence that reading from a computer screen is slower than from a book [244]. This is thought to be

due to a number of factors including a longer line length, fewer words to a page,

1.2 Input–output channels **23**

orientation and the familiarity of the medium of the page. These factors can of course be reduced by careful design of textual interfaces.

A final word about the use of contrast in visual display: a negative contrast (dark characters on a light screen) provides higher luminance and, therefore, increased acuity, than a positive contrast. This will in turn increase legibility. However, it will also be more prone to flicker. Experimental evidence suggests that in practice negat ive contrast displays are preferred and result in more accurate performance [30].

1.2.2 Hearing

The sense of hearing is often considered secondary to sight, but we tend to under estimate the amount of information that we receive through our ears. Close your eyes for a moment and listen. What sounds can you hear? Where are they coming from? What is making them? As I sit at my desk I can hear cars passing on the road outside, machinery working on a site nearby, the drone of a plane overhead and bird song. But I can also tell *where* the sounds are coming from, and estimate how far away they are. So from the sounds I hear I can tell that a car is passing on a particular road near my house, and which direction it is traveling in. I know that building work is in progress in a particular location, and that a certain type of bird is perched in the tree in my garden.

The auditory system can convey a lot of information about our environment. But how does it work?

*The human ear*

Just as vision begins with light, hearing begins with vibrations in the air or *sound waves*. The ear receives these vibrations and transmits them, through various stages, to the auditory nerves. The ear comprises three sections, commonly known as the *outer ear*, *middle ear* and *inner ear*.

The outer ear is the visible part of the ear. It has two parts: the *pinna*, which is the structure that is attached to the sides of the head, and the *auditory canal*, along which sound waves are passed to the middle ear. The outer ear serves two purposes. First, it protects the sensitive middle ear from damage. The auditory canal contains wax which prevents dust, dirt and over-inquisitive insects reaching the middle ear. It also maintains the middle ear at a constant temperature. Secondly, the pinna and auditory canal serve to amplify some sounds.

The middle ear is a small cavity connected to the outer ear by the *tympanic membrane*, or ear drum, and to the inner ear by the *cochlea*. Within the cavity are the *ossicles*, the smallest bones in the body. Sound waves pass along the auditory canal and vibrate the ear drum which in turn vibrates the ossicles, which transmit the vibrations to the cochlea, and so into the inner ear. This ‘relay’ is required because, unlike the air-filled outer and middle ears, the inner ear is filled with a denser cochlean liquid. If passed directly from the air to the liquid, the transmission of the sound waves would be poor. By transmitting them via the ossicles the sound waves are concentrated and amplified.

**24** Chapter 1 n The human

The waves are passed into the liquid-filled cochlea in the inner ear. Within the cochlea are delicate hair cells or *cilia* that bend because of the vibrations in the cochlean liquid and release a chemical transmitter which causes impulses in the auditory nerve.

*Processing sound*

As we have seen, sound is changes or vibrations in air pressure. It has a number of characteristics which we can differentiate. *Pitch* is the frequency of the sound. A low frequency produces a low pitch, a high frequency, a high pitch. *Loudness* is propor tional to the amplitude of the sound; the frequency remains constant. *Timbre* relates to the type of the sound: sounds may have the same pitch and loudness but be made by different instruments and so vary in timbre. We can also identify a sound’s loca tion, since the two ears receive slightly different sounds, owing to the time difference between the sound reaching the two ears and the reduction in intensity caused by the sound waves reflecting from the head.

The human ear can hear frequencies from about 20 Hz to 15 kHz. It can distin guish frequency changes of less than 1.5 Hz at low frequencies but is less accurate at high frequencies. Different frequencies trigger activity in neurons in different parts of the auditory system, and cause different rates of firing of nerve impulses.

The auditory system performs some filtering of the sounds received, allowing us to ignore background noise and concentrate on important information. We are selective in our hearing, as illustrated by the *cocktail party effect*, where we can pick out our name spoken across a crowded noisy room. However, if sounds are too loud, or frequencies too similar, we are unable to differentiate sound.

As we have seen, sound can convey a remarkable amount of information. It is rarely used to its potential in interface design, usually being confined to warning sounds and notifications. The exception is multimedia, which may include music, voice commentary and sound effects. However, the ear can differentiate quite subtle sound changes and can recognize familiar sounds without concentrating attention on the sound source. This suggests that sound could be used more extensively in interface design, to convey information about the system state, for example. This is discussed in more detail in Chapter 10.

Worked exercise *Suggest ideas for an interface which uses the properties of sound effectively.*

Answer You might approach this exercise by considering how sound could be added to an appli cation with which you are familiar. Use your imagination. This is also a good subject for a literature survey (starting with the references in Chapter 10).

Speech sounds can obviously be used to convey information. This is useful not only for the visually impaired but also for any application where the user’s attention has to be divided (for example, power plant control, flight control, etc.). Uses of non-speech sounds include the following:

n Attention – to attract the user’s attention to a critical situation or to the end of a process, for example.

1.2 Input–output channels **25**

n Status information – continuous background sounds can be used to convey status information. For example, monitoring the progress of a process (without the need for visual attention).

n Confirmation – a sound associated with an action to confirm that the action has been carried out. For example, associating a sound with deleting a file. n Navigation – using changing sound to indicate where the user is in a system. For example, what about sound to support navigation in hypertext?

1.2.3 Touch

The third and last of the senses that we will consider is touch or *haptic perception*. Although this sense is often viewed as less important than sight or hearing, imagine life without it. Touch provides us with vital information about our environment. It tells us when we touch something hot or cold, and can therefore act as a warning. It also provides us with feedback when we attempt to lift an object, for example. Con sider the act of picking up a glass of water. If we could only see the glass and not feel when our hand made contact with it or feel its shape, the speed and accuracy of the action would be reduced. This is the experience of users of certain *virtual reality* games: they can see the computer-generated objects which they need to manipulate but they have no physical sensation of touching them. Watching such users can be an informative and amusing experience! Touch is therefore an important means of feedback, and this is no less so in using computer systems. Feeling buttons depress is an important part of the task of pressing the button. Also, we should be aware that, although for the average person, haptic perception is a secondary source of informa tion, for those whose other senses are impaired, it may be vitally important. For such users, interfaces such as braille may be the primary source of information in the interaction. We should not therefore underestimate the importance of touch.

The apparatus of touch differs from that of sight and hearing in that it is not local ized. We receive stimuli through the skin. The skin contains three types of sensory receptor: *thermoreceptors* respond to heat and cold, *nociceptors* respond to intense pressure, heat and pain, and *mechanoreceptors* respond to pressure. It is the last of these that we are concerned with in relation to human–computer interaction.

There are two kinds of mechanoreceptor, which respond to different types of pressure. *Rapidly adapting mechanoreceptors* respond to immediate pressure as the skin is indented. These receptors also react more quickly with increased pressure. However, they stop responding if continuous pressure is applied. *Slowly adapting mechanoreceptors* respond to continuously applied pressure.

Although the whole of the body contains such receptors, some areas have greater sensitivity or acuity than others. It is possible to measure the acuity of different areas of the body using the *two-point threshold test*. Take two pencils, held so their tips are about 12 mm apart. Touch the points to your thumb and see if you can feel two points. If you cannot, move the points a little further apart. When you can feel two points, measure the distance between them. The greater the distance, the lower the sensitivity. You can repeat this test on different parts of your body. You should find

**26** Chapter 1 n The human

that the measure on the forearm is around 10 times that of the finger or thumb. The fingers and thumbs have the highest acuity.

A second aspect of haptic perception is *kinesthesis*: awareness of the position of the body and limbs. This is due to receptors in the joints. Again there are three types: rapidly adapting, which respond when a limb is moved in a particular direc tion; slowly adapting, which respond to both movement and static position; and positional receptors, which only respond when a limb is in a static position. This perception affects both comfort and performance. For example, for a touch typist, awareness of the relative positions of the fingers and feedback from the keyboard are very important.

Handling the goods E-commerce has become very successful in some areas of sales, such as travel services, books and CDs, and food. However, in some retail areas, such as clothes shopping, e-commerce has been less successful. Why?

When buying train and airline tickets and, to some extent, books and food, the experience of shop ping is less important than the convenience. So, as long as we know what we want, we are happy to shop online. With clothes, the experience of shopping is far more important. We need to be able to handle the goods, feel the texture of the material, check the weight to test quality. Even if we know that something will fit us we still want to be able to handle it before buying.

Research into haptic interaction (see Chapter 2 and Chapter 10) is looking at ways of solving this problem. By using special force feedback and tactile hardware, users are able to feel surfaces and shape. For example, a demonstration environment called TouchCity allows people to walk around a virtual shopping mall, pick up products and feel their texture and weight. A key problem with the commercial use of such an application, however, is that the haptic experience requires expensive hardware not yet available to the average e-shopper. However, in future, such immer sive e-commerce experiences are likely to be the norm. (See www.novint.com/)

1.2.4 Movement

Before leaving this section on the human’s input–output channels, we need to consider motor control and how the way we move affects our interaction with com puters. A simple action such as hitting a button in response to a question involves a number of processing stages. The stimulus (of the question) is received through the sensory receptors and transmitted to the brain. The question is processed and a valid response generated. The brain then tells the appropriate muscles to respond. Each of these stages takes time, which can be roughly divided into reaction time and movement time.

Movement time is dependent largely on the physical characteristics of the subjects: their age and fitness, for example. Reaction time varies according to the sensory channel through which the stimulus is received. A person can react to an auditory

1.3 Human memory **27**

signal in approximately 150 ms, to a visual signal in 200 ms and to pain in 700 ms. However, a combined signal will result in the quickest response. Factors such as skill or practice can reduce reaction time, and fatigue can increase it.

A second measure of motor skill is accuracy. One question that we should ask is whether speed of reaction results in reduced accuracy. This is dependent on the task and the user. In some cases, requiring increased reaction time reduces accuracy. This is the premise behind many arcade and video games where less skilled users fail at levels of play that require faster responses. However, for skilled operators this is not necessarily the case. Studies of keyboard operators have shown that, although the faster operators were up to twice as fast as the others, the slower ones made 10 times the errors.

Speed and accuracy of movement are important considerations in the design of interactive systems, primarily in terms of the time taken to move to a particular target on a screen. The target may be a button, a menu item or an icon, for example. The time taken to hit a target is a function of the size of the target and the distance that has to be moved. This is formalized in *Fitts’ law* [135]. There are many vari ations of this formula, which have varying constants, but they are all very similar. One common form is

Movement time = *a* + *b* log2(distance/size + 1)

where *a* and *b* are empirically determined constants.

This affects the type of target we design. Since users will find it more difficult to manipulate small objects, targets should generally be as large as possible and the distance to be moved as small as possible. This has led to suggestions that pie chart-shaped menus are preferable to lists since all options are equidistant. However, the trade-off is increased use of screen estate, so the choice may not be so simple. If lists are used, the most frequently used options can be placed closest to the user’s start point (for example, at the top of the menu). The implications of Fitts’ law in design are discussed in more detail in Chapter 12.

**1.3**

HUMAN MEMORY

Have you ever played the memory game? The idea is that each player has to recount a list of objects and add one more to the end. There are many variations but the objects are all loosely related: ‘I went to the market and bought a lemon, some oranges, bacon . . .’ or ‘I went to the zoo and saw monkeys, and lions, and tigers . . .’ and so on. As the list grows objects are missed out or recalled in the wrong order and so people are eliminated from the game. The winner is the person remaining at the end. Such games rely on our ability to store and retrieve information, even seemingly arbitrary items. This is the job of our memory system.

Indeed, much of our everyday activity relies on memory. As well as storing all our factual knowledge, our memory contains our knowledge of actions or procedures.

**28** Chapter 1 n The human

**Figure 1.9** A model of the structure of memory

It allows us to repeat actions, to use language, and to use new information received via our senses. It also gives us our sense of identity, by preserving information from our past experiences.

But how does our memory work? How do we remember arbitrary lists such as those generated in the memory game? Why do some people remember more easily than others? And what happens when we forget?

In order to answer questions such as these, we need to understand some of the capabilities and limitations of human memory. Memory is the second part of our model of the human as an information-processing system. However, as we noted earlier, such a division is simplistic since, as we shall see, memory is associated with each level of processing. Bearing this in mind, we will consider the way in which memory is structured and the activities that take place within the system.

It is generally agreed that there are three types of memory or memory function: *sensory buffers*, *short-term memory* or *working memory*, and *long-term memory*. There is some disagreement as to whether these are three separate systems or different functions of the same system. We will not concern ourselves here with the details

of this debate, which is discussed in detail by Baddeley [21], but will indicate the evidence used by both sides as we go along. For our purposes, it is sufficient to note three separate types of memory. These memories interact, with information being processed and passed between memory stores, as shown in Figure 1.9.

1.3.1 Sensory memory

The sensory memories act as buffers for stimuli received through the senses. A sensory memory exists for each sensory channel: *iconic memory* for visual stimuli, *echoic memory* for aural stimuli and *haptic memory* for touch. These memories are constantly overwritten by new information coming in on these channels.

We can demonstrate the existence of iconic memory by moving a finger in front of the eye. Can you see it in more than one place at once? This indicates a persistence of the image after the stimulus has been removed. A similar effect is noticed most vividly at firework displays where moving sparklers leave a persistent image. Information remains in iconic memory very briefly, in the order of 0.5 seconds.

Similarly, the existence of echoic memory is evidenced by our ability to ascertain the direction from which a sound originates. This is due to information being received by both ears. However, since this information is received at different times, we must store the stimulus in the meantime. Echoic memory allows brief ‘play-back’

1.3 Human memory **29**

of information. Have you ever had someone ask you a question when you are reading? You ask them to repeat the question, only to realize that you know what was asked after all. This experience, too, is evidence of the existence of echoic memory.

Information is passed from sensory memory into short-term memory by atten tion, thereby filtering the stimuli to only those which are of interest at a given time. Attention is the concentration of the mind on one out of a number of competing stimuli or thoughts. It is clear that we are able to focus our attention selectively, choosing to attend to one thing rather than another. This is due to the limited capa city of our sensory and mental processes. If we did not selectively attend to the

stimuli coming into our senses, we would be overloaded. We can choose which stimuli to attend to, and this choice is governed to an extent by our *arousal*, our level of interest or need. This explains the cocktail party phenomenon mentioned earlier: we can attend to one conversation over the background noise, but we may choose

to switch our attention to a conversation across the room if we hear our name mentioned. Information received by sensory memories is quickly passed into a more permanent memory store, or overwritten and lost.

1.3.2 Short-term memory

Short-term memory or working memory acts as a ‘scratch-pad’ for temporary recall of information. It is used to store information which is only required fleetingly. For example, calculate the multiplication 35 ⋅ 6 in your head. The chances are that you will have done this calculation in stages, perhaps 5 ⋅ 6 and then 30 ⋅ 6 and added

the results; or you may have used the fact that 6 = 2 ⋅ 3 and calculated 2 ⋅ 35 = 70 followed by 3 ⋅ 70. To perform calculations such as this we need to store the inter mediate stages for use later. Or consider reading. In order to comprehend this sentence you need to hold in your mind the beginning of the sentence as you read the rest. Both of these tasks use short-term memory.

Short-term memory can be accessed rapidly, in the order of 70 ms. However, it also decays rapidly, meaning that information can only be held there temporarily, in the order of 200 ms.

Short-term memory also has a limited capacity. There are two basic methods for measuring memory capacity. The first involves determining the length of a sequence which can be remembered in order. The second allows items to be freely recalled in any order. Using the first measure, the average person can remember 7 ± 2 digits. This was established in experiments by Miller [234]. Try it. Look at the following number sequence:

265397620853

Now write down as much of the sequence as you can remember. Did you get it all right? If not, how many digits could you remember? If you remembered between five and nine digits your *digit span* is average.

Now try the following sequence:

44 113 245 8920

**30** Chapter 1 n The human

Did you recall that more easily? Here the digits are grouped or *chunked*. A general ization of the 7 ± 2 rule is that we can remember 7 ± 2 *chunks* of information. Therefore chunking information can increase the short-term memory capacity. The limited capacity of short-term memory produces a subconscious desire to create chunks, and so optimize the use of the memory. The successful formation of a chunk is known as *closure*. This process can be generalized to account for the desire to com plete or close tasks held in short-term memory. If a subject fails to do this or is pre vented from doing so by interference, the subject is liable to lose track of what she is doing and make consequent errors.

DESIGN FOCUS 

Cashing in

Closure gives you a nice ‘done it’ when we complete some part of a task. At this point our minds have a tendency to flush short-term memory in order to get on with the next job. Early automatic teller machines (ATMs) gave the customer money before returning their bank card. On receiving the money the customer would reach closure and hence often forget to take the card. Modern ATMs return the card first!

The sequence of chunks given above also makes use of pattern abstraction: it is written in the form of a UK telephone number which makes it easier to remember. We may even recognize the first sets of digits as the international code for the UK and the dialing code for Leeds – chunks of information. Patterns can be useful as aids

1.3 Human memory **31**

to memory. For example, most people would have difficulty remembering the fol lowing sequence of chunks:

HEC ATR ANU PTH ETR EET

However, if you notice that by moving the last character to the first position, you get the statement ‘the cat ran up the tree’, the sequence is easy to recall. In experiments where subjects were able to recall words freely, evidence shows that recall of the last words presented is better than recall of those in the middle [296]. This is known as the *recency effect*. However, if the subject is asked to perform another task between presentation and recall (for example, counting backwards) the recency effect is eliminated. The recall of the other words is unaffected. This suggests that short-term memory recall is damaged by interference of other information. However, the fact that this interference does not affect recall of earlier items provides some evidence for the existence of separate long-term and short-term memories. The early items are held in a long-term store which is unaffected by the recency effect. Interference does not necessarily impair recall in short-term memory. Baddeley asked subjects to remember six-digit numbers and attend to sentence processing at the same time [21]. They were asked to answer questions on sentences, such as ‘A precedes B: AB is true or false?’. Surprisingly, this did not result in interference, suggesting that in fact short-term memory is not a unitary system but is made up of a number of components, including a visual channel and an articulatory channel. The task of sen tence processing used the visual channel, while the task of remembering digits used the articulatory channel, so interference only occurs if tasks utilize the same channel. These findings led Baddeley to propose a model of working memory that incorp orated a number of elements together with a central processing executive. This is illustrated in Figure 1.10.

**Figure 1.10** A more detailed model of short-term memory

**32** Chapter 1 n The human

DESIGN FOCUS 

7 ± 2 revisited

When we looked at short-term memory, we noted the general rule that people can hold 7 ± 2 items or chunks of information in short-term memory. It is a principle that people tend to remember but it can be misapplied. For example, it is often suggested that this means that lists, menus and other groups of items should be designed to be no more than 7 items long. But use of menus and lists of course has little to do with short-term memory – they are available in the environment as cues and so do not need to be remembered.

On the other hand the 7 ± 2 rule would apply in command line interfaces. Imagine a scenario where a UNIX user looks up a command in the manual. Perhaps the command has a number of parameters of options, to be applied in a particular order, and it is going to be applied to several files that have long path names. The user then has to hold the command, its parameters and the file path names in short term memory while he types them in. Here we could say that the task may cause problems if the num ber of items or chunks in the command line string is more than 7.

1.3.3 Long-term memory

If short-term memory is our working memory or ‘scratch-pad’, long-term memory is our main resource. Here we store factual information, experiential knowledge, procedural rules of behavior – in fact, everything that we ‘know’. It differs from short-term memory in a number of significant ways. First, it has a huge, if not unlim ited, capacity. Secondly, it has a relatively slow access time of approximately a tenth of a second. Thirdly, forgetting occurs more slowly in long-term memory, if at all. These distinctions provide further evidence of a memory structure with several parts.

Long-term memory is intended for the long-term storage of information. Information is placed there from working memory through rehearsal. Unlike work ing memory there is little decay: long-term recall after minutes is the same as that after hours or days.

*Long-term memory structure*

There are two types of long-term memory: *episodic memory* and *semantic memory*. Episodic memory represents our memory of events and experiences in a serial form. It is from this memory that we can reconstruct the actual events that took place at a given point in our lives. Semantic memory, on the other hand, is a structured record of facts, concepts and skills that we have acquired. The information in semantic memory is derived from that in our episodic memory, such that we can learn new facts or concepts from our experiences.

Semantic memory is structured in some way to allow access to information, representation of relationships between pieces of information, and inference. One model for the way in which semantic memory is structured is as a network. Items are

1.3 Human memory **33**

**Figure 1.11** Long-term memory may store information in a semantic network

associated to each other in classes, and may inherit attributes from parent classes. This model is known as a *semantic network*. As an example, our knowledge about dogs may be stored in a network such as that shown in Figure 1.11.

Specific breed attributes may be stored with each given breed, yet general dog information is stored at a higher level. This allows us to generalize about specific cases. For instance, we may not have been told that the sheepdog Shadow has four legs and a tail, but we can infer this information from our general knowledge about sheepdogs and dogs in general. Note also that there are connections within the net work which link into other domains of knowledge, for example cartoon characters. This illustrates how our knowledge is organized by association.

The viability of semantic networks as a model of memory organization has been demonstrated by Collins and Quillian [74]. Subjects were asked questions about different properties of related objects and their reaction times were measured. The types of question asked (taking examples from our own network) were ‘Can a collie breathe?’, ‘Is a beagle a hound?’ and ‘Does a hound track?’ In spite of the fact that the answers to such questions may seem obvious, subjects took longer to answer ques tions such as ‘Can a collie breathe?’ than ones such as ‘Does a hound track?’ The

reason for this, it is suggested, is that in the former case subjects had to search fur ther through the memory hierarchy to find the answer, since information is stored at its most abstract level.

A number of other memory structures have been proposed to explain how we represent and store different types of knowledge. Each of these represents a different

**34** Chapter 1 n The human

**Figure 1.12** A frame-based representation of knowledge

aspect of knowledge and, as such, the models can be viewed as complementary rather than mutually exclusive. Semantic networks represent the associations and relation ships between single items in memory. However, they do not allow us to model the representation of more complex objects or events, which are perhaps composed of

a number of items or activities. Structured representations such as *frames* and *scripts* organize information into data structures. *Slots* in these structures allow attribute values to be added. Frame slots may contain default, fixed or variable information. A frame is instantiated when the slots are filled with appropriate values. Frames and scripts can be linked together in networks to represent hierarchical structured knowledge.

Returning to the ‘dog’ domain, a frame-based representation of the knowledge may look something like Figure 1.12. The fixed slots are those for which the attribute value is set, default slots represent the usual attribute value, although this may be overridden in particular instantiations (for example, the Basenji does not bark), and variable slots can be filled with particular values in a given instance. Slots can also contain procedural knowledge. Actions or operations can be associated with a slot and performed, for example, whenever the value of the slot is changed.

Frames extend semantic nets to include structured, hierarchical information. They represent knowledge items in a way which makes explicit the relative importance of each piece of information.

Scripts attempt to model the representation of stereotypical knowledge about situ ations. Consider the following sentence:

John took his dog to the surgery. After seeing the vet, he left.

From our knowledge of the activities of dog owners and vets, we may fill in a substantial amount of detail. The animal was ill. The vet examined and treated the animal. John paid for the treatment before leaving. We are less likely to assume the alternative reading of the sentence, that John took an instant dislike to the vet on sight and did not stay long enough to talk to him!

**Figure 1.13** A script for visiting the vet

1.3 Human memory **35**

A script represents this default or stereotypical information, allowing us to inter pret partial descriptions or cues fully. A script comprises a number of elements, which, like slots, can be filled with appropriate information:

**Entry conditions** Conditions that must be satisfied for the script to be activated. **Result** Conditions that will be true after the script is terminated. **Props** Objects involved in the events described in the script.

**Roles** Actions performed by particular participants.

**Scenes** The sequences of events that occur.

**Tracks** A variation on the general pattern representing an alternative scenario.

An example script for going to the vet is shown in Figure 1.13.

A final type of knowledge representation which we hold in memory is the repre sentation of procedural knowledge, our knowledge of how to do something. A com mon model for this is the production system. Condition–action rules are stored in long-term memory. Information coming into short-term memory can match a condition in one of these rules and result in the action being executed. For example, a pair of production rules might be

IF dog is wagging tail

THEN pat dog

IF dog is growling

THEN run away

If we then meet a growling dog, the condition in the second rule is matched, and we respond by turning tail and running. (Not to be recommended by the way!)

**36** Chapter 1 n The human

*Long-term memory processes*

So much for the structure of memory, but what about the processes which it uses? There are three main activities related to long-term memory: storage or remember ing of information, forgetting and information retrieval. We shall consider each of these in turn.

First, how does information get into long-term memory and how can we improve this process? Information from short-term memory is stored in long-term memory by rehearsal. The repeated exposure to a stimulus or the rehearsal of a piece of informa tion transfers it into long-term memory.

This process can be optimized in a number of ways. Ebbinghaus performed numerous experiments on memory, using himself as a subject [117]. In these experi ments he tested his ability to learn and repeat nonsense syllables, comparing his recall minutes, hours and days after the learning process. He discovered that the amount learned was directly proportional to the amount of time spent learning.

This is known as the *total time hypothesis*. However, experiments by Baddeley and others suggest that learning time is most effective if it is distributed over time [22]. For example, in an experiment in which Post Office workers were taught to type, those whose training period was divided into weekly sessions of one hour performed better than those who spent two or four hours a week learning (although the former obviously took more weeks to complete their training). This is known as the *distribu tion of practice effect*.

However, repetition is not enough to learn information well. If information is not meaningful it is more difficult to remember. This is illustrated by the fact that it is more difficult to remember a set of words representing concepts than a set of words representing objects. Try it. First try to remember the words in list A and test yourself.

List A: Faith Age Cold Tenet Quiet Logic Idea Value Past Large

Now try list B.

List B: Boat Tree Cat Child Rug Plate Church Gun Flame Head

The second list was probably easier to remember than the first since you could visualize the objects in the second list.

Sentences are easier still to memorize. Bartlett performed experiments on remem bering meaningful information (as opposed to meaningless such as Ebbinghaus used) [28]. In one such experiment he got subjects to learn a story about an un familiar culture and then retell it. He found that subjects would retell the story replacing unfamiliar words and concepts with words which were meaningful to them. Stories were effectively translated into the subject’s own culture. This is related to the semantic structuring of long-term memory: if information is meaningful and familiar, it can be related to existing structures and more easily incorporated into memory.

Memorable or secure?

1.3 Human memory **37 **

As online activities become more widespread, people are having to remember more and more access information, such as passwords and security checks. The average active internet user may have separate passwords and user names for several email accounts, mailing lists, e-shopping sites, e-banking, online auctions and more! Remembering these passwords is not easy.

From a security perspective it is important that passwords are random. Words and names are very easy to crack, hence the recommendation that passwords are frequently changed and constructed from random strings of letters and numbers. But in reality these are the hardest things for people to commit to memory. Hence many people will use the same password for all their online activities (rarely if ever changing it) and will choose a word or a name that is easy for them to remember, in spite of the obviously increased security risks. Security here is in conflict with memorability!

A solution to this is to construct a nonsense password out of letters or numbers that will have meaning to you but will not make up a word in a dictionary (e.g. initials of names, numbers from significant dates or postcodes, and so on). Then what is remembered is the meaningful rule for constructing the password, and not a meaningless string of alphanumeric characters.

So if structure, familiarity and concreteness help us in learning information, what causes us to lose this information, to forget? There are two main theories of forget ting: *decay* and *interference*. The first theory suggests that the information held in long-term memory may eventually be forgotten. Ebbinghaus concluded from his experiments with nonsense syllables that information in memory decayed logarith mically, that is that it was lost rapidly to begin with, and then more slowly. *Jost’s law*, which follows from this, states that if two memory traces are equally strong at a given time the older one will be more durable.

The second theory is that information is lost from memory through interference. If we acquire new information it causes the loss of old information. This is termed *retroactive interference*. A common example of this is the fact that if you change tele phone numbers, learning your new number makes it more difficult to remember your old number. This is because the new association masks the old. However, some times the old memory trace breaks through and interferes with new information. This is called *proactive inhibition*. An example of this is when you find yourself driv ing to your old house rather than your new one.

Forgetting is also affected by emotional factors. In experiments, subjects given emotive words and non-emotive words found the former harder to remember in the short term but easier in the long term. Indeed, this observation tallies with our experience of selective memory. We tend to remember positive information rather than negative (hence nostalgia for the ‘good old days’), and highly emotive events rather than mundane.

**38** Chapter 1 n The human

It is debatable whether we ever actually forget anything or whether it just becomes increasingly difficult to access certain items from memory. This question is in some ways moot since it is impossible to prove that we *do* forget: appearing to have for gotten something may just be caused by not being able to retrieve it! However, there is evidence to suggest that we may not lose information completely from long-term memory. First, proactive inhibition demonstrates the recovery of old information even after it has been ‘lost’ by interference. Secondly, there is the ‘tip of the tongue’ experience, which indicates that some information is present but cannot be satisfac torily accessed. Thirdly, information may not be recalled but may be recognized, or may be recalled only with prompting.

This leads us to the third process of memory: information retrieval. Here we need to distinguish between two types of information retrieval, recall and recognition. In recall the information is reproduced from memory. In recognition, the presentation of the information provides the knowledge that the information has been seen before. Recognition is the less complex cognitive activity since the information is provided as a cue.

However, recall can be assisted by the provision of retrieval cues, which enable the subject quickly to access the information in memory. One such cue is the use of categories. In an experiment subjects were asked to recall lists of words, some of which were organized into categories and some of which were randomly organized.

The words that were related to a category were easier to recall than the others [38]. Recall is even more successful if subjects are allowed to categorize their own lists of words during learning. For example, consider the following list of words:

child red plane dog friend blood cold tree big angry

Now make up a story that links the words using as vivid imagery as possible. Now try to recall as many of the words as you can. Did you find this easier than the previous experiment where the words were unrelated?

The use of vivid imagery is a common cue to help people remember information. It is known that people often visualize a scene that is described to them. They can then answer questions based on their visualization. Indeed, subjects given a descrip tion of a scene often embellish it with additional information. Consider the follow ing description and imagine the scene:

The engines roared above the noise of the crowd. Even in the blistering heat people rose to their feet and waved their hands in excitement. The flag fell and they were off. Within seconds the car had pulled away from the pack and was careering round the bend at a desperate pace. Its wheels momentarily left the ground as it cornered. Coming down the straight the sun glinted on its shimmering paint. The driver gripped the wheel with fierce concentration. Sweat lay in fine drops on his brow.

Without looking back to the passage, what color is the car?

If you could answer that question you have visualized the scene, including the car’s color. In fact, the color of the car is not mentioned in the description at all.

Improve your memory

1.4 Thinking: reasoning and problem solving **39 **

Many people can perform astonishing feats of memory: recalling the sequence of cards in a pack (or multiple packs – up to six have been reported), or recounting π to 1000 decimal places, for example. There are also adverts to ‘Improve Your Memory’ (usually leading to success, or wealth, or other such inducement), and so the question arises: can you improve your memory abilities? The answer is yes; this exercise shows you one technique.

Look at the list below of numbers and associated words:

1 bun 6 sticks

2 shoe 7 heaven

3 tree 8 gate

4 door 9 wine

5 hive 10 hen

Notice that the words sound similar to the numbers. Now think about the words one at a time and visualize them, in as much detail as possible. For example, for ‘1’, think of a large, sticky iced bun, the base spiralling round and round, with raisins in it, covered in sweet, white, gooey icing. Now do the rest, using as much visualization as you can muster: imagine how things would look, smell, taste, sound, and so on.

This is your reference list, and you need to know it off by heart.

Having learnt it, look at a pile of at least a dozen odd items collected together by a colleague. The task is to look at the collection of objects for only 30 seconds, and then list as many as possible without making a mistake or viewing the collection again. Most people can manage between five and eight items, if they do not know any memory-enhancing techniques like the following.

Mentally pick one (say, for example, a paper clip), and call it number one. Now visualize it inter acting with the bun. It can get stuck into the icing on the top of the bun, and make your fingers all gooey and sticky when you try to remove it. If you ate the bun without noticing, you’d get a crunched tooth when you bit into it – imagine how that would feel. When you’ve really got a graphic scenario developed, move on to the next item, call it number two, and again visualize it interacting with the reference item, shoe. Continue down your list, until you have done 10 things.

This should take you about the 30 seconds allowed. Then hide the collection and try and recall the numbers in order, the associated reference word, and then the image associated with that word. You should find that you can recall the 10 associated items practically every time. The technique can be easily extended by extending your reference list.

**1.4**

THINKING: REASONING AND PROBLEM SOLVING

We have considered how information finds its way into and out of the human system and how it is stored. Finally, we come to look at how it is processed and manipulated. This is perhaps the area which is most complex and which separates

**40** Chapter 1 n The human

humans from other information-processing systems, both artificial and natural. Although it is clear that animals receive and store information, there is little evid ence to suggest that they can use it in quite the same way as humans. Similarly, artificial intelligence has produced machines which can see (albeit in a limited way) and store information. But their ability to use that information is limited to small domains.

Humans, on the other hand, are able to use information to reason and solve problems, and indeed do these activities when the information is partial or unavail able. Human thought is conscious and self-aware: while we may not always be able to identify the processes we use, we can identify the products of these processes, our thoughts. In addition, we are able to think about things of which we have no experience, and solve problems which we have never seen before. How is this done?

Thinking can require different amounts of knowledge. Some thinking activities are very directed and the knowledge required is constrained. Others require vast amounts of knowledge from different domains. For example, performing a subtrac tion calculation requires a relatively small amount of knowledge, from a constrained domain, whereas understanding newspaper headlines demands knowledge of pol itics, social structures, public figures and world events.

In this section we will consider two categories of thinking: reasoning and problem solving. In practice these are not distinct since the activity of solving a problem may well involve reasoning and vice versa. However, the distinction is a common one and is helpful in clarifying the processes involved.

1.4.1 Reasoning

*Reasoning* is the process by which we use the knowledge we have to draw conclusions or infer something new about the domain of interest. There are a number of differ ent types of reasoning: *deductive*, *inductive* and *abductive*. We use each of these types of reasoning in everyday life, but they differ in significant ways.

*Deductive reasoning*

Deductive reasoning derives the logically necessary conclusion from the given pre mises. For example,

If it is Friday then she will go to work

It is Friday

Therefore she will go to work.

It is important to note that this is the *logical* conclusion from the premises; it does not necessarily have to correspond to our notion of truth. So, for example,

If it is raining then the ground is dry

It is raining

Therefore the ground is dry.

1.4 Thinking: reasoning and problem solving **41**

is a perfectly valid deduction, even though it conflicts with our knowledge of what is true in the world.

Deductive reasoning is therefore often misapplied. Given the premises

Some people are babies

Some babies cry

many people will infer that ‘Some people cry’. This is in fact an invalid deduction since we are not told that all babies are people. It is therefore logically possible that the babies who cry are those who are not people.

It is at this point, where truth and validity clash, that human deduction is poorest. One explanation for this is that people bring their world knowledge into the reason ing process. There is good reason for this. It allows us to take short cuts which make dialog and interaction between people informative but efficient. We assume a certain amount of shared knowledge in our dealings with each other, which in turn allows us to interpret the inferences and deductions implied by others. If validity rather than truth was preferred, all premises would have to be made explicit.

*Inductive reasoning*

Induction is generalizing from cases we have seen to infer information about cases we have not seen. For example, if every elephant we have ever seen has a trunk, we infer that all elephants have trunks. Of course, this inference is unreliable and cannot be proved to be true; it can only be proved to be false. We can disprove the inference simply by producing an elephant without a trunk. However, we can never prove it true because, no matter how many elephants with trunks we have seen or are known to exist, the next one we see may be trunkless. The best that we can do is gather evid ence to support our inductive inference.

In spite of its unreliability, induction is a useful process, which we use constantly in learning about our environment. We can never see all the elephants that have ever lived or will ever live, but we have certain knowledge about elephants which we are prepared to trust for all practical purposes, which has largely been inferred by induc tion. Even if we saw an elephant without a trunk, we would be unlikely to move from our position that ‘All elephants have trunks’, since we are better at using positive than negative evidence. This is illustrated in an experiment first devised by Wason [365]. You are presented with four cards as in Figure 1.14. Each card has a number on one side and a letter on the other. Which cards would you need to pick up to test the truth of the statement ‘If a card has a vowel on one side it has an even number on the other’?

A common response to this (was it yours?) is to check the E and the 4. However, this uses only positive evidence. In fact, to test the truth of the statement we need to check negative evidence: if we can find a card which has an odd number on one side and a vowel on the other we have disproved the statement. We must therefore check E and 7. (It does not matter what is on the other side of the other cards: the state ment does not say that all even numbers have vowels, just that all vowels have even numbers.)

**42** Chapter 1 n The human

**Figure 1.14** Wason’s cards

Filling the gaps Look again at Wason’s cards in Figure 1.14. In the text we say that you only need to check the E and the 7. This is correct, but only because we very carefully stated in the text that ‘each card has a number on one side and a letter on the other’. If the problem were stated without that condition then the K would also need to be examined in case it has a vowel on the other side. In fact, when the problem is so stated, even the most careful subjects ignore this possibility. Why? Because the nature of the problem implicitly suggests that each card has a number on one side and a letter on the other.

This is similar to the embellishment of the story at the end of Section 1.3.3. In fact, we constantly fill in gaps in the evidence that reaches us through our senses. Although this can lead to errors in our reasoning it is also essential for us to function. In the real world we rarely have all the evid ence necessary for logical deductions and at all levels of perception and reasoning we fill in details in order to allow higher levels of reasoning to work.

*Abductive reasoning*

The third type of reasoning is abduction. Abduction reasons from a fact to the action or state that caused it. This is the method we use to derive explanations for the events we observe. For example, suppose we know that Sam always drives too fast when she has been drinking. If we see Sam driving too fast we may infer that she has been drinking. Of course, this too is unreliable since there may be another reason why she is driving fast: she may have been called to an emergency, for example.

In spite of its unreliability, it is clear that people do infer explanations in this way, and hold onto them until they have evidence to support an alternative theory or explanation. This can lead to problems in using interactive systems. If an event always follows an action, the user will infer that the event is caused by the action unless evidence to the contrary is made available. If, in fact, the event and the action are unrelated, confusion and even error often result.

1.4 Thinking: reasoning and problem solving **43**

1.4.2 Problem solving

If reasoning is a means of inferring new information from what is already known, problem solving is the process of finding a solution to an unfamiliar task, using the knowledge we have. Human problem solving is characterized by the ability to adapt the information we have to deal with new situations. However, often solutions seem to be original and creative. There are a number of different views of how people

solve problems. The earliest, dating back to the first half of the twentieth century, is the *Gestalt* view that problem solving involves both reuse of knowledge and insight. This has been largely superseded but the questions it was trying to address remain and its influence can be seen in later research. A second major theory, proposed in the 1970s by Newell and Simon, was the *problem space theory*, which takes the view that the mind is a limited information processor. Later variations on this drew on the earlier theory and attempted to reinterpret Gestalt theory in terms of information processing theories. We will look briefly at each of these views.

*Gestalt theory*

Gestalt psychologists were answering the claim, made by behaviorists, that prob lem solving is a matter of reproducing known responses or trial and error. This explanation was considered by the Gestalt school to be insufficient to account for human problem-solving behavior. Instead, they claimed, problem solving is both *pro ductive* and *reproductive*. Reproductive problem solving draws on previous experi ence as the behaviorists claimed, but productive problem solving involves insight and restructuring of the problem. Indeed, reproductive problem solving could be a hind rance to finding a solution, since a person may ‘fixate’ on the known aspects of the problem and so be unable to see novel interpretations that might lead to a solution.

Gestalt psychologists backed up their claims with experimental evidence. Kohler provided evidence of apparent insight being demonstrated by apes, which he observed joining sticks together in order to reach food outside their cages [202]. However, this was difficult to verify since the apes had once been wild and so could have been using previous knowledge.

Other experiments observed human problem-solving behavior. One well-known example of this is Maier’s *pendulum problem* [224]. The problem was this: the subjects were in a room with two pieces of string hanging from the ceiling. Also in the room were other objects including pliers, poles and extensions. The task set was to tie the pieces of string together. However, they were too far apart to catch hold of both at once. Although various solutions were proposed by subjects, few chose to use the weight of the pliers as a pendulum to ‘swing’ the strings together. How ever, when the experimenter brushed against the string, setting it in motion, this solution presented itself to subjects. Maier interpreted this as an example of produc tive restructuring. The movement of the string had given insight and allowed the subjects to see the problem in a new way. The experiment also illustrates fixation: subjects were initially unable to see beyond their view of the role or use of a pair of pliers.

**44** Chapter 1 n The human

Although Gestalt theory is attractive in terms of its description of human problem solving, it does not provide sufficient evidence or structure to support its theories. It does not explain when restructuring occurs or what insight is, for example. How ever, the move away from behaviorist theories was helpful in paving the way for the information-processing theory that was to follow.

*Problem space theory*

Newell and Simon proposed that problem solving centers on the problem space. The problem space comprises *problem states*, and problem solving involves generating these states using legal state transition operators. The problem has an initial state and a goal state and people use the operators to move from the former to the latter. Such problem spaces may be huge, and so *heuristics* are employed to select appro priate operators to reach the goal. One such heuristic is *means–ends analysis*. In means–ends analysis the initial state is compared with the goal state and an oper ator chosen to reduce the difference between the two. For example, imagine you are reorganizing your office and you want to move your desk from the north wall of the room to the window. Your initial state is that the desk is at the north wall. The goal state is that the desk is by the window. The main difference between these two is the location of your desk. You have a number of operators which you can apply to mov ing things: you can carry them or push them or drag them, etc. However, you know that to carry something it must be light and that your desk is heavy. You therefore have a new subgoal: to make the desk light. Your operators for this may involve removing drawers, and so on.

An important feature of Newell and Simon’s model is that it operates within the constraints of the human processing system, and so searching the problem space is limited by the capacity of short-term memory, and the speed at which information can be retrieved. Within the problem space framework, experience allows us to solve problems more easily since we can structure the problem space appropriately and choose operators efficiently.

Newell and Simon’s theory, and their *General Problem Solver* model which is based on it, have largely been applied to problem solving in well-defined domains, for example solving puzzles. These problems may be unfamiliar but the knowledge that is required to solve them is present in the statement of the problem and the expected solution is clear. In real-world problems finding the knowledge required to solve the problem may be part of the problem, or specifying the goal may be difficult. Problems such as these require significant domain knowledge: for example, to solve a programming problem you need knowledge of the language and the domain in which the program operates. In this instance specifying the goal clearly may be a significant part of solving the problem.

However, the problem space framework provides a clear theory of problem solving, which can be extended, as we shall see when we look at skill acquisition in the next section, to deal with knowledge-intensive problem solving. First we will look briefly at the use of analogy in problem solving.

1.4 Thinking: reasoning and problem solving **45**

Worked exercise *Identify the goals and operators involved in the problem ‘delete the second paragraph of the document’ on a word processor. Now use a word processor to delete a paragraph and note your actions, goals and subgoals. How well did they match your earlier description?*

Answer Assume you have a document open and you are at some arbitrary position within it. You also need to decide which operators are available and what their preconditions and results are. Based on an imaginary word processor we assume the following operators (you may wish to use your own WP package):

Operator Precondition Result

delete\_paragraph Cursor at start of paragraph Paragraph deleted

move\_to\_paragraph Cursor anywhere in document Cursor moves to start of

next paragraph (except

where there is no next

paragraph when no effect)

move\_to\_start Cursor anywhere in document Cursor at start of document

**Goal**: *delete second paragraph in document*

Looking at the operators an obvious one to resolve this goal is delete\_paragraph which has the precondition ‘cursor at start of paragraph’. We therefore have a new subgoal: *move\_to\_paragraph*. The precondition is ‘cursor anywhere in document’ (which we can meet) but we want the second paragraph so we must initially be in the first.

We set up a new subgoal, move\_to\_start, with precondition ‘cursor anywhere in docu ment’ and result ‘cursor at start of document’. We can then apply *move\_to\_paragraph* and finally *delete\_paragraph*.

We assume some knowledge here (that the second paragraph is the paragraph after the first one).

*Analogy in problem solving*

A third element of problem solving is the use of analogy. Here we are interested in how people solve novel problems. One suggestion is that this is done by mapping knowledge relating to a similar known domain to the new problem – called *analo gical mapping*. Similarities between the known domain and the new one are noted and operators from the known domain are transferred to the new one.

This process has been investigated using analogous stories. Gick and Holyoak [149] gave subjects the following problem:

A doctor is treating a malignant tumor. In order to destroy it he needs to blast it with high-intensity rays. However, these will also destroy the healthy tissue sur rounding the tumor. If he lessens the rays’ intensity the tumor will remain. How does he destroy the tumor?

**46** Chapter 1 n The human

The solution to this problem is to fire low-intensity rays from different directions converging on the tumor. That way, the healthy tissue receives harmless low intensity rays while the tumor receives the rays combined, making a high-intensity dose. The investigators found that only 10% of subjects reached this solution with out help. However, this rose to 80% when they were given this analogous story and told that it may help them:

A general is attacking a fortress. He can’t send all his men in together as the roads are mined to explode if large numbers of men cross them. He therefore splits his men into small groups and sends them in on separate roads.

In spite of this, it seems that people often miss analogous information, unless it is semantically close to the problem domain. When subjects were not told to use the story, many failed to see the analogy. However, the number spotting the analogy rose when the story was made semantically close to the problem, for example a general using rays to destroy a castle.

The use of analogy is reminiscent of the Gestalt view of productive restructuring and insight. Old knowledge is used to solve a new problem.

1.4.3 Skill acquisition

All of the problem solving that we have considered so far has concentrated on handling unfamiliar problems. However, for much of the time, the problems that we face are not completely new. Instead, we gradually acquire skill in a particular domain area. But how is such skill acquired and what difference does it make to our problem-solving performance? We can gain insight into how skilled behavior works, and how skills are acquired, by considering the difference between novice and expert behavior in given domains.

Chess: of human and artificial intelligence A few years ago, Deep Blue, a chess-playing computer, beat Gary Kasparov, the world’s top Grand Master, in a full tournament. This was the long-awaited breakthrough for the artificial intelligence (AI) community, who have traditionally seen chess as the ultimate test of their art. However, despite the fact that computer chess programs can play at Grand Master level against human players, this does not mean they play in the same way. For each move played, Deep Blue investigated many millions of alternative moves and counter-moves. In contrast, a human chess player will only consider a few dozen. But, if the human player is good, these will usually be the right few dozen. The ability to spot patterns allows a human to address a problem with far less effort than a brute force approach. In chess, the number of moves is such that finally brute force, applied fast enough, has overcome human pattern-matching skill. In Go, which has far more pos sible moves, computer programs do not even reach a good club level of play. Many models of the mental processes have been heavily influenced by computation. It is worth remembering that although there are similarities, computer ‘intelligence’ is very different from that of humans.

1.4 Thinking: reasoning and problem solving **47**

A commonly studied domain is chess playing. It is particularly suitable since it lends itself easily to representation in terms of problem space theory. The initial state is the opening board position; the goal state is one player checkmating the other; operators to move states are legal moves of chess. It is therefore possible to examine skilled behavior within the context of the problem space theory of problem solving.

Studies of chess players by DeGroot, Chase and Simon, among others, produced some interesting observations [64, 65, 88, 89]. In all the experiments the behavior of chess masters was compared with less experienced chess players. The first observa tion was that players did not consider large numbers of moves in choosing their move, nor did they look ahead more than six moves (often far fewer). Masters con sidered no more alternatives than the less experienced, but they took less time to make a decision and produced better moves.

So what makes the difference between skilled and less skilled behavior in chess? It appears that chess masters remember board configurations and good moves associated with them. When given actual board positions to remember, masters are much better at reconstructing the board than the less experienced. However,

when given random configurations (which were unfamiliar), the groups of players were equally bad at reconstructing the positions. It seems therefore that expert players ‘chunk’ the board configuration in order to hold it in short-term memory. Expert players use larger chunks than the less experienced and can therefore re member more detail.

This behavior is also seen among skilled computer programmers. They can also reconstruct programs more effectively than novices since they have the structures available to build appropriate chunks. They acquire plans representing code to solve particular problems. When that problem is encountered in a new domain or new program they will recall that particular plan and reuse it.

Another observed difference between skilled and less skilled problem solving is in the way that different problems are grouped. Novices tend to group problems according to superficial characteristics such as the objects or features common to both. Experts, on the other hand, demonstrate a deeper understanding of the prob lems and group them according to underlying conceptual similarities which may not be at all obvious from the problem descriptions.

Each of these differences stems from a better encoding of knowledge in the expert: information structures are fine tuned at a deep level to enable efficient and accurate retrieval. But how does this happen? How is skill such as this acquired? One model of skill acquisition is Anderson’s *ACT\** model [14]. ACT\* identifies three basic

levels of skill:

1. The learner uses general-purpose rules which interpret facts about a problem. This is slow and demanding on memory access.

2. The learner develops rules specific to the task.

3. The rules are tuned to speed up performance.

General mechanisms are provided to account for the transitions between these levels. For example, *proceduralization* is a mechanism to move from the first to the second. It removes the parts of the rule which demand memory access and replaces

**48** Chapter 1 n The human

variables with specific values. *Generalization*, on the other hand, is a mechanism which moves from the second level to the third. It generalizes from the specific cases to general properties of those cases. Commonalities between rules are condensed to produce a general-purpose rule.

These are best illustrated by example. Imagine you are learning to cook. Initially you may have a general rule to tell you how long a dish needs to be in the oven, and a number of explicit representations of dishes in memory. You can instantiate the rule by retrieving information from memory.

IF cook[type, ingredients, time]

THEN

cook for: time

cook[casserole, [chicken,carrots,potatoes], 2 hours]

cook[casserole, [beef,dumplings,carrots], 2 hours]

cook[cake, [flour,sugar,butter,eggs], 45 mins]

Gradually your knowledge becomes proceduralized and you have specific rules for each case:

IF type is casserole

AND ingredients are [chicken,carrots,potatoes]

THEN

cook for: 2 hours

IF type is casserole

AND ingredients are [beef,dumplings,carrots]

THEN

cook for: 2 hours

IF type is cake

AND ingredients are [flour,sugar,butter,eggs]

THEN

cook for: 45 mins

Finally, you may generalize from these rules to produce general-purpose rules, which exploit their commonalities:

IF type is casserole

AND ingredients are ANYTHING

THEN

cook for: 2 hours

The first stage uses knowledge extensively. The second stage relies upon known procedures. The third stage represents skilled behavior. Such behavior may in fact become automatic and as such be difficult to make explicit. For example, think of an activity at which you are skilled, perhaps driving a car or riding a bike. Try to describe to someone the exact procedure which you go through to do this. You will find this quite difficult. In fact experts tend to have to rehearse their actions mentally in order to identify exactly what they do. Such skilled behavior is efficient but may cause errors when the context of the activity changes.

1.4 Thinking: reasoning and problem solving **49**

1.4.4 Errors and mental models

Human capability for interpreting and manipulating information is quite impres sive. However, we do make mistakes. Some are trivial, resulting in no more than temporary inconvenience or annoyance. Others may be more serious, requiring substantial effort to correct. Occasionally an error may have catastrophic effects, as we see when ‘human error’ results in a plane crash or nuclear plant leak.

Why do we make mistakes and can we avoid them? In order to answer the latter part of the question we must first look at what is going on when we make an error. There are several different types of error. As we saw in the last section some errors result from changes in the context of skilled behavior. If a pattern of behavior has become automatic and we change some aspect of it, the more familiar pattern may break through and cause an error. A familiar example of this is where we intend to stop at the shop on the way home from work but in fact drive past. Here, the activ ity of driving home is the more familiar and overrides the less familiar intention.

Other errors result from an incorrect understanding, or model, of a situation or system. People build their own theories to understand the causal behavior of sys tems. These have been termed *mental models*. They have a number of characteristics. Mental models are often partial: the person does not have a full understanding of the working of the whole system. They are unstable and are subject to change. They can be internally inconsistent, since the person may not have worked through the logical consequences of their beliefs. They are often unscientific and may be based on super stition rather than evidence. Often they are based on an incorrect interpretation of the evidence.

DESIGN FOCUS 

Human error and false memories

In the second edition of this book, one of the authors added the following story:

During the Second World War a new cockpit design was introduced for Spitfires. The pilots were trained and flew successfully during training, but would unaccountably bail out when engaged in dog fights. The new design had exchanged the positions of the gun trigger and ejector controls. In the heat of battle the old responses resurfaced and the pilots ejected. Human error, yes, but the designer’s error, not the pilot’s.

It is a good story, but after the book was published we got several emails saying ‘Spitfires didn’t have ejector seats’. It was Kai-Mikael Jää-Aro who was able to find what may have been the original to the story (and incidentally inform us what model of Spitfire was in our photo and who the pilot was!). He pointed us to and translated the story of Sierra 44, an S35E Draken reconnaissance aircraft.1 The full story involves just about every perceptual and cognitive error imaginable, but the point that links to

1. Pej Kristoffersson, 1984. Sigurd 44 – Historien om hur man gör bort sig så att det märks by, *Flygrevyn* 2/1984, pp. 44–6.

**50** Chapter 1 n The human

the (false) Spitfire story is that in the Draken the red buttons for releasing the fuel ‘drop’ tanks and for the canopy release differed only in very small writing. In an emergency (burning fuel tanks) the pilot accidentally released the canopy and so ended up flying home cabriolet style.

There is a second story of human error here – the author’s memory. When the book was written he could not recall where he had come across the story but was convinced it was to do with a Spitfire. It may be that he had been told the story by someone else who had got it mixed up, but it is as likely that he simply remembered the rough outline of the story and then ‘reconstructed’ the rest. In fact that is exactly how our memories work. Our brains do not bother to lay down every little detail, but when we ‘remember’ we rebuild what the incident ‘must have been’ using our world knowledge. This pro cess is completely unconscious and can lead to what are known as *false memories*. This is particularly problematic in witness statements in criminal trials as early questioning by police or lawyers can unin tentionally lead to witnesses being sure they have seen things that they have not. Numerous controlled psychological experiments have demonstrated this effect which furthermore is strongly influenced by biasing factors such as the race of supposed criminals.

To save his blushes we have not said here which author’s failing memory was responsible for the Spitfire story, but you can read more on this story and also find who it was on the book website at: /e3/online/spitfire/

Courtesy of popperfoto.com

Assuming a person builds a mental model of the system being dealt with, errors may occur if the actual operation differs from the mental model. For example, on one occasion we were staying in a hotel in Germany, attending a conference. In the lobby of the hotel was a lift. Beside the lift door was a button. Our model of the sys tem, based on previous experience of lifts, was that the button would call the lift. We pressed the button and the lobby light went out! In fact the button was a light switch and the lift button was on the inside rim of the lift, hidden from view.

1.5 Emotion **51**

Although both the light switch and the lift button were inconsistent with our men tal models of these controls, we would probably have managed if they had been encountered separately. If there had been no button beside the lift we would have looked more closely and found the one on the inner rim. But since the light switch reflected our model of a lift button we looked no further. During our stay we observed many more new guests making the same error.

This illustrates the importance of a correct mental model and the dangers of ignoring conventions. There are certain conventions that we use to interpret the world and ideally designs should support these. If these are to be violated, explicit support must be given to enable us to form a correct mental model. A label on the button saying ‘light switch’ would have been sufficient.

**1.5**

EMOTION

So far in this chapter we have concentrated on human perceptual and cognitive abil ities. But human experience is far more complex than this. Our emotional response to situations affects how we perform. For example, positive emotions enable us to think more creatively, to solve complex problems, whereas negative emotion pushes us into narrow, focussed thinking. A problem that may be easy to solve when we are relaxed, will become difficult if we are frustrated or afraid.

Psychologists have studied emotional response for decades and there are many theories as to what is happening when we feel an emotion and why such a response occurs. More than a century ago, William James proposed what has become known as the James–Lange theory (Lange was a contemporary of James whose theories

were similar): that emotion was the interpretation of a physiological response, rather than the other way around. So while we may feel that we respond *to* an emotion, James contended that we respond physiologically to a stimulus and interpret that as emotion:

Common sense says, we lose our fortune, are sorry and weep; we meet a bear, are frightened and run; we are insulted by a rival, are angry and strike. The hypothesis here . . . is that we feel sorry because we cry, angry because we strike, afraid because we tremble.

(W. James, *Principles of Psychology*, page 449. Henry Holt, New York, 1890.)

Others, however, disagree. Cannon [54a], for example, argued that our physio logical processes are in fact too slow to account for our emotional reactions, and that the physiological responses for some emotional states are too similar (e.g. anger and fear), yet they can be easily distinguished. Experience in studies with the use of drugs that stimulate broadly the same physiological responses as anger or fear seems to support this: participants reported physical symptoms but not the emotion, which suggests that emotional response is more than a recognition of physiological changes.

**52** Chapter 1 n The human

Schachter and Singer [312a] proposed a third interpretation: that emotion results from a person evaluating physical responses in the light of the whole situation. So whereas the same physiological response can result from a range of different situ ations, the emotion that is felt is based on a cognitive evaluation of the circumstance and will depend on what the person attributes this to. So the same physiological response of a pounding heart will be interpreted as excitement if we are in a com petition and fear if we find ourselves under attack.

Whatever the exact process, what is clear is that emotion involves both physical and cognitive events. Our body responds biologically to an external stimulus and we interpret that in some way as a particular emotion. That biological response – known as *affect* – changes the way we deal with different situations, and this has an impact on the way we interact with computer systems. As Donald Norman says:

Negative affect can make it harder to do even easy tasks; positive affect can make it easier to do difficult tasks.

(D. A. Norman, Emotion and design: attractive things work better.

*Interactions Magazine*, ix(4): 36–42, 2002.)

So what are the implications of this for design? It suggests that in situations of stress, people will be less able to cope with complex problem solving or managing difficult interfaces, whereas if people are relaxed they will be more forgiving of limitations in the design. This does not give us an excuse to design bad interfaces but does suggest that if we build interfaces that promote positive responses – for

example by using aesthetics or reward – then they are likely be more successful.

**1.6**

INDIVIDUAL DIFFERENCES

In this chapter we have been discussing humans in general. We have made the assumption that everyone has similar capabilities and limitations and that we can therefore make generalizations. To an extent this is true: the psychological principles and properties that we have discussed apply to the majority of people. Notwithstanding this, we should remember that, although we share processes in common, humans, and therefore users, are not all the same. We should be aware of individual differences so that we can account for them as far as possible within our designs. These differences may be long term, such as sex, physical capabilities and intellectual capabilities. Others are shorter term and include the effect of stress or fatigue on the user. Still others change through time, such as age.

These differences should be taken into account in our designs. It is useful to consider, for any design decision, if there are likely to be users within the target group who will be adversely affected by our decision. At the extremes a decision may exclude a section of the user population. For example, the current emphasis on visual interfaces excludes those who are visually impaired, unless the design also makes use of the other sensory channels. On a more mundane level, designs should allow for

1.7 Psychology and the design of interactive systems **53**

users who are under pressure, feeling ill or distracted by other concerns: they should not push users to their perceptual or cognitive limits.

We will consider the issues of universal accessibility in more detail in Chapter 10.

**1.7**

PSYCHOLOGY AND THE DESIGN OF

INTERACTIVE SYSTEMS

So far we have looked briefly at the way in which humans receive, process and store information, solve problems and acquire skill. But how can we apply what we have learned to designing interactive systems? Sometimes, straightforward conclu sions can be drawn. For example, we can deduce that recognition is easier than recall and allow users to select commands from a set (such as a menu) rather than input them directly. However, in the majority of cases, application is not so obvious or simple. In fact, it may be dangerous, leading us to make generalizations which are not valid. In order to apply a psychological principle or result properly in design, we need to understand its context, both in terms of where it fits in the wider field of psychology and in terms of the details of the actual experiments, the measures used and the subjects involved, for example. This may appear daunting, particularly to the novice designer who wants to acknowledge the relevance of cognitive psy chology but does not have the background to derive appropriate conclusions. Fortunately, principles and results from research in psychology have been distilled into guidelines for design, models to support design and techniques for evaluating design. Parts 2 and 3 of this book include discussion of a range of guidelines, models and techniques, based on cognitive psychology, which can be used to support the design process.

1.7.1 Guidelines

Throughout this chapter we have discussed the strengths and weaknesses of human cognitive and perceptual processes but, for the most part, we have avoided attempt ing to apply these directly to design. This is because such an attempt could only be partial and simplistic, and may give the impression that this is all psychology has to offer.

However, general design principles and guidelines can be and have been derived from the theories we have discussed. Some of these are relatively straightforward: for instance, recall is assisted by the provision of retrieval cues so interfaces should incorporate recognizable cues wherever possible. Others are more complex and con text dependent. In Chapter 7 we discuss principles and guidelines further, many of which are derived from psychological theory. The interested reader is also referred to Gardiner and Christie [140] which illustrates how guidelines can be derived from psychological theory.