Formalizing Commonsense Psychology: Chapter Summaries

A Comprehensive Analysis of Gordon & Hobbs 2017

Chapter Summary Analysis

2025

Contents

This document contains comprehensive summaries and analyses of all 49 chapters from "Formalizing Commonsense Psychology" by Gordon & Hobbs (2017). Each chapter summary includes detailed analysis across nine key dimensions:

1. **Key Features Identified** - Core concepts, axioms, and theoretical elements
2. **Technical Sophistication** - Complexity of formal machinery and logical frameworks

3. **Complexity Distribution** - Categorization of axioms by difficulty level
4. **Conceptual Importance** - Theoretical significance and foundational role
5. **Cross-Chapter Connections** - Relationships and dependencies with other chapters
6. **Applications Mentioned** - Practical applications and use cases
7. **Notable Design Decisions** - Key architectural and methodological choices
8. **Theoretical Significance** - Broader implications for commonsense reasoning
9. **Unique Contributions** - Novel innovations and distinctive achievements

The book presents a comprehensive formalization of commonsense psychology through 892 axioms across 45 theories, representing the most ambitious attempt to systematically encode human psychological reasoning in formal logic.

Organization: - **Chapters 1-4**: Methodological foundations and linguistic analysis
- **Chapters 5-20**: Background theories (mathematical and logical foundations)
- **Chapters 21-49**: Commonsense psychology theories (knowledge, memory, planning, emotions, etc.)

Chapter 1

Chapter 1: Commonsense Psychology and Psychology

1.1 Overview

- **Introduction to commonsense psychology** through Heider & Simmel's famous triangle experiments (1944)
- **Historical development** from Gestalt psychology to modern Theory of Mind research
- **Cross-disciplinary analysis** spanning developmental psychology, autism research, cross-cultural studies, primate cognition, and philosophy of mind
- **Contemporary debates** between scientific psychology and folk psychology, including eliminative materialism

1.2 Key Features Identified:

1.2.1 1. Heider-Simmel Triangle Experiment Foundation:

- **Anthropomorphic Attribution**: 114 Smith College students spontaneously attributed mental states (beliefs, desires, emotions) to geometric shapes
- **Narrative Construction**: Subjects created coherent stories involving intentional agents, conflicts, and psychological motivations
- **Theory of Mind Demonstration**: Even simple visual trajectories trigger rich psychological explanations when physics-based explanations fail

1.2.2 2. Heider's Commonsense Psychology Framework:

- **Theoretical Systematization**: First formal attempt to describe folk psychology as coherent system of beliefs

- **Logical Notation**: Created symbolic logic notation for interpersonal relations (p, o, q for persons; x, y, z for entities)
- **Ten Central Concepts**: Subjective environment, perceiving, causing, can, trying, wanting, suffering, sentiments, belonging, ought
- **Predictive Power**: Focus on commonsense psychology's utility for everyday interpersonal prediction and explanation

1.2.3 3. Theory of Mind Research Evolution:

- **Developmental Shifts**: False-belief task performance shows systematic development from ages 3-5
- **Theory Theory vs. Simulation Theory**: Major theoretical divide over whether ToM involves theoretical knowledge or mental simulation
- **Autism Connection**: Deficits in false-belief performance linked to autism spectrum disorders
- **Cross-Cultural Universality**: Core abilities appear across cultures with some variation in specific beliefs

1.2.4 4. Historical Change in Commonsense Psychology:

- **Unconscious Concepts**: Evidence for cultural shifts in psychological concepts (e.g., Freudian unconscious becoming commonplace)
- **Literary Analysis**: Study of English novels (1813-1922) shows introduction of unconscious desire concepts around 1850s
- **Cultural Transmission**: Psychological theories can migrate from scientific to folk psychology through cultural mechanisms

1.2.5 5. Scientific vs. Folk Psychology Tensions:

- **Predictive Limitations**: Folk psychology fails to predict many experimental psychology findings (e.g., Langer effect)
- **Gap Persistence**: Limited evidence for successful transfer from scientific to folk psychology
- **Eliminative Materialism**: Radical proposal to replace folk psychology with neuroscience-based theories

1.3 Technical Sophistication:

- **Interdisciplinary Integration**: Synthesizes findings from developmental psychology, cognitive science, anthropology, and philosophy
- **Methodological Diversity**: Experimental studies, cross-cultural research, historical analysis, and theoretical argumentation
- **Theoretical Frameworks**: Multiple competing theories (Theory Theory, Simulation Theory, eliminative materialism)

- **Empirical Grounding**: Extensive experimental evidence from false-belief tasks, autism studies, and cross-species comparisons

1.4 Complexity Distribution:

- **Historical Foundation**: 30% devoted to Heider's original framework and Attribution Theory development
- **Contemporary Research**: 50% covering developmental psychology, autism, cross-cultural, and primate studies
- **Philosophical Analysis**: 20% examining eliminative materialism and science-folk psychology relationships

1.5 Conceptual Importance:

1.5.1 Foundational Significance:

Establishes commonsense psychology as fundamental human capacity that enables social interaction, communication, and cultural transmission. Shows how geometric shapes can trigger sophisticated psychological attributions, demonstrating the pervasiveness and automaticity of folk psychological reasoning.

1.5.2 Scientific Framework:

Provides theoretical foundation for understanding relationship between scientific and folk psychology. Argues against simple replacement models, suggesting instead that scientific psychology extends and refines commonsense understanding rather than replacing it wholesale.

1.5.3 Interdisciplinary Bridge:

Connects multiple research traditions by showing how commonsense psychology research spans developmental psychology, autism studies, cross-cultural research, primate cognition, and philosophy of mind.

1.6 Cross-Chapter Connections:

- **Chapters 2-4**: Introduces themes continued throughout Part I regarding relationship between commonsense psychology and formal systems
- **Parts II-III**: Provides psychological foundation for formal logical system developed in subsequent parts
- **Entire Work**: Establishes fundamental motivation for formalizing commonsense psychology theories

1.7 Applications Mentioned:

- **Experimental Psychology**: False-belief tasks, autism diagnosis, developmental assessment
- **Cross-Cultural Research**: Understanding cultural variations in folk psychology concepts
- **Artificial Intelligence**: Need for machines to understand human folk psychology for interaction
- **Education**: Science education challenges in changing folk psychological beliefs
- **Philosophy of Mind**: Eliminative materialism debates about future of folk psychology

1.8 Notable Design Decisions:

1.8.1 Historical Grounding:

Begins with Heider's foundational work to establish continuity between early attribution research and contemporary Theory of Mind studies, showing how questions persist across decades of research.

1.8.2 Cross-Disciplinary Synthesis:

Integrates findings from multiple fields rather than focusing on single research tradition, demonstrating breadth of commonsense psychology research and its interdisciplinary relevance.

1.8.3 Balanced Perspective:

Presents multiple theoretical positions (Theory Theory, Simulation Theory, eliminative materialism) without strongly advocating for particular view, allowing reader to appreciate complexity of debates.

1.8.4 Cultural Evolution Framework:

Emphasizes that commonsense psychology can change historically, providing foundation for understanding how formal theories might influence folk beliefs.

1.8.5 Practical Focus:

Emphasizes functional utility of folk psychology for everyday social interaction, arguing against eliminative approaches based on practical necessity.

1.9 Theoretical Significance:

Chapter 1 establishes commonsense psychology as both a fundamental human capacity and a legitimate object of scientific study. The chapter's strength lies in showing how simple experimental paradigms (geometric shapes moving on screen) can reveal sophisticated psychological reasoning processes that operate automatically and universally across cultures.

The historical analysis demonstrates that commonsense psychology is not static but can evolve through cultural transmission of scientific ideas. This provides crucial foundation for the book's larger project of formalizing commonsense psychology - such formalization could potentially influence folk psychology if successfully transmitted through culture.

The interdisciplinary synthesis shows how commonsense psychology research connects developmental psychology, autism studies, cross-cultural research, primate cognition, and philosophy of mind. This breadth demonstrates the fundamental importance of folk psychological reasoning for human cognition and social interaction.

The chapter's treatment of eliminative materialism provides important counterpoint to overly optimistic views about replacing folk psychology with scientific theories. The argument that commonsense psychology is "extraordinarily powerful" and "remarkably easy for human beings to use" establishes pragmatic foundation for descriptive rather than eliminative approaches.

1.10 Unique Contributions:

1.10.1 Triangle Film Analysis:

Detailed analysis of Heider-Simmel experiments shows how minimal visual stimuli can trigger maximal psychological attribution, demonstrating both the power and automaticity of folk psychological reasoning.

1.10.2 Historical Linguistics Analysis:

Novel computational analysis of English novels (1813-1922) provides empirical evidence for historical change in commonsense psychology concepts, specifically the introduction of unconscious desire concepts in mid-19th century.

1.10.3 Eliminative Materialism Critique:

Provides sustained philosophical argument against eliminative approaches based on practical necessity and cultural entrenchment of folk psychological concepts in language, law, and social institutions.

1.10.4 Science-Folk Psychology Gap:

Systematic analysis of why scientific psychology findings fail to transfer to folk psychology, arguing for extension rather than replacement model of scientific influence.

1.10.5 Cross-Species Framework:

Comparative analysis of Theory of Mind abilities across species shows both continuity (goal understanding) and discontinuity (false-belief reasoning) between humans and other primates.

This chapter establishes the intellectual foundation for the book's formal approach to commonsense psychology by demonstrating both its fundamental importance for human social life and its susceptibility to systematic scientific analysis and potential cultural modification.

Chapter 2

Chapter 2: Commonsense Psychology and Computers

2.1 Overview

- **Anthropomorphism as human universal** - examining how people naturally attribute human psychology to non-human entities
- **Human-computer interaction challenges** - analyzing gaps between user expectations and computational behavior
- **Desktop metaphor critique** - arguing that current interfaces discourage natural anthropomorphic tendencies
- **Anthropomorphic computing proposal** - advocating for systems that accommodate user anthropomorphism through commonsense psychology models

2.2 Key Features Identified:

2.2.1 1. Universal Anthropomorphism Framework:

- **Evolutionary Utility**: Anthropomorphism as adaptive strategy for managing complexity in environment
- **Development Continuum**: From childhood toy attribution to adult institutional personification (governments, corporations)
- **Legal Recognition**: US Supreme Court cases establishing corporate personhood and free speech rights
- **Dennett's Intentional Stance**: Strategic attribution of beliefs, desires, and rational agency for behavioral prediction

2.2.2 2. Human-Computer Discord Sources:

- **Attention Focus Mismatch**: Users expect single-task focus while computers efficiently multthread dozens of processes
- **User Model Poverty**: Computers lack understanding of user mental states, goals, and contextual needs
- **Psychological Language Gaps**: Limited natural language understanding of mental state concepts despite their centrality to human communication
- **Counterintuitive Operations**: System behavior violates commonsense psychological expectations about intelligent behavior

2.2.3 3. Desktop Metaphor Analysis:

- **Historical Context**: Developed at Xerox PARC (1970s), popularized by Apple and Microsoft (1980s) for office worker market
- **Anthropomorphism Barrier**: Office objects are lifeless and familiar, reducing tendency to see computers as people
- **Turing Test Connection**: Early computing vision (Star Trek, HAL 9000) assumed anthropomorphic interaction
- **Future Obsolescence**: Ubiquitous computing will require more natural interaction metaphors than office environments

2.2.4 4. Anthropomorphic Computing Vision:

- **Postage-Stamp Computers**: Future scenario with ubiquitous, cheap computing requiring natural interaction
- **Natural User Instruction**: "You're a toaster" - direct anthropomorphic assignment of roles and capabilities
- **Commonsense Psychology Alignment**: Systems must operate according to how people think intelligent devices should behave
- **Facilitated Anthropomorphism**: Rather than discouraging, systems should support and leverage user anthropomorphic tendencies

2.2.5 5. Computational Modeling Distinctions:

- **Cognitive Models vs. Inferential Theories**: Models simulate reasoning processes vs. theories describe knowledge content
- **Knowledge Engineering Requirements**: Large-scale effort needed, not simple algorithmic solutions
- **Content Theory Approach**: Focus on axiom lists defining commonsense consequences rather than syntactic complexity

- **AI Engineering Methodology**: Systematic formalization approach for computational implementation

2.3 Technical Sophistication:

2.3.1 Interdisciplinary Integration:

- **Psychology**: Dennett's intentional stance, Reeves & Nass media equation research
- **Computer Science**: Human-computer interaction, knowledge representation, AI reasoning systems
- **Philosophy**: Eliminative materialism connections, theory of mind debates
- **Legal Studies**: Corporate personhood cases and institutional anthropomorphism

2.3.2 Human-Computer Interaction Analysis:

- **Media Equation Research**: Experimental evidence that people treat computers as social agents
- **Politeness Effects**: Users inflate performance judgments when computers directly ask for feedback
- **Social Role Attribution**: Automatic application of human social behaviors to computational systems
- **Design Implications**: Alignment of user social/cognitive expectations with system behavior

2.3.3 Engineering Challenges:

- **Natural Language Processing**: Understanding psychological vocabulary and mental state references
- **User Modeling**: Tracking beliefs, goals, plans, preferences, attention states throughout interactions
- **Contextual Reasoning**: Dynamic computation of implications based on changing user mental models
- **Explanation Generation**: Producing comprehensible statements about system states and processes

2.4 Complexity Distribution:

- **Anthropomorphism Theory**: 25% covering universality, development, and strategic utility
- **Human-Computer Discord**: 30% analyzing specific interaction problems and their psychological foundations
- **Interface Design Critique**: 25% examining desktop metaphor limitations and future computing visions

- **Implementation Framework**: 20% outlining computational requirements and engineering approaches

2.5 Conceptual Importance:

2.5.1 Technological Transformation:

Demonstrates how computing is transitioning from office-based tools to ubiquitous environmental presence, requiring fundamental changes in interaction paradigms. Shows that anthropomorphism is inevitable rather than optional in human-technology relationships.

2.5.2 Psychological Realism:

Argues that effective human-computer interaction must accommodate rather than fight natural human psychological tendencies. Provides evidence that people automatically apply social and intentional reasoning to computational systems regardless of design intentions.

2.5.3 Engineering Requirements:

Establishes that anthropomorphic computing requires explicit computational models of commonsense psychology rather than simple interface improvements. Demonstrates need for large-scale knowledge engineering efforts in psychological domain.

2.6 Cross-Chapter Connections:

- **Chapter 1**: Builds on commonsense psychology foundation by showing its necessity for technological interaction
- **Chapters 3-4**: Motivates need for formal logical approaches to representing psychological knowledge
- **Parts II-III**: Provides practical justification for detailed formalization work in subsequent chapters

2.7 Applications Mentioned:

- **Word Processing**: Analysis of current applications' psychological blindness to user mental states
- **Future Ubiquitous Computing**: Postage-stamp computers requiring natural anthropomorphic interaction
- **Gaming and Multimedia**: Examples where anthropomorphic expectations create user frustration
- **Corporate and Government Systems**: Large-scale anthropomorphization in organizational contexts

- **Voice Assistants and Smart Devices**: Implied applications for current anthropomorphic computing

2.8 Notable Design Decisions:

2.8.1 Anti-Desktop Argument:

Provides sustained critique of dominant interface paradigm, arguing it artificially suppresses natural anthropomorphic tendencies that should be embraced rather than discouraged.

2.8.2 Future Computing Vision:

Uses concrete scenarios (postage-stamp computers) to illustrate how ubiquitous computing will require anthropomorphic interaction methods, grounding abstract arguments in tangible examples.

2.8.3 Dennett Integration:

Adopts intentional stance as theoretical framework for understanding anthropomorphism, providing cognitive science foundation for engineering recommendations.

2.8.4 Discord Analysis:

Systematic examination of specific ways current systems violate psychological expectations, providing concrete requirements for anthropomorphic system design.

2.8.5 Engineering Realism:

Acknowledges that anthropomorphic computing requires substantial engineering effort rather than simple algorithmic solutions, setting realistic expectations for implementation challenges.

2.9 Theoretical Significance:

Chapter 2 establishes anthropomorphic computing as both inevitable and beneficial for human-technology interaction. The chapter's strength lies in showing how anthropomorphism emerges from fundamental human cognitive architecture rather than from naive beliefs about technology.

The systematic analysis of human-computer discord demonstrates that current interaction problems stem from mismatches between human psychological expectations and system behavior. This analysis provides specific requirements for anthropomorphic system design: attention modeling, user mental state tracking, and psychologically-grounded natural language processing.

The historical analysis of the desktop metaphor shows how interface design can artificially suppress natural human tendencies. The argument that ubiquitous computing will make anthropomorphic interaction necessary provides strong motivation for developing the required computational models.

The chapter's treatment of computational modeling distinguishes between cognitive process simulation and knowledge content representation, clarifying the specific type of psychological modeling needed for anthropomorphic computing. This distinction is crucial for understanding the book's subsequent formal approach.

2.10 Unique Contributions:

2.10.1 Anthropomorphic Computing Framework:

First systematic argument for designing computational systems to facilitate rather than discourage user anthropomorphism, providing specific requirements for implementation.

2.10.2 Desktop Metaphor Critique:

Historical analysis showing how current interface paradigms artificially suppress natural human psychological tendencies, arguing for fundamental design philosophy changes.

2.10.3 Future Computing Vision:

Concrete scenarios demonstrating how ubiquitous computing will require anthropomorphic interaction methods, grounding abstract arguments in tangible technological evolution.

2.10.4 Discord Taxonomy:

Systematic classification of specific ways current systems violate psychological expectations: attention focus, user modeling, and language understanding gaps.

2.10.5 Engineering Requirements Analysis:

Realistic assessment of computational challenges for anthropomorphic computing, acknowledging need for large-scale knowledge engineering rather than simple algorithmic solutions.

2.10.6 Smedslund Integration:

Detailed analysis of existing psychological formalization work (Psychologic), identifying strengths and limitations for computational implementation.

This chapter establishes the practical necessity for the formal commonsense psychology work that follows in subsequent chapters. It demonstrates that anthropomorphic computing is not merely desirable but inevitable given trends in computing ubiquity and human cognitive architecture. The chapter provides both theoretical foundation and practical motivation for the detailed formalization work that comprises the remainder of the book.

Chapter 3

Chapter 3: Formalizing Commonsense Psychology

3.1 Key Features Identified:

3.1.1 1. Successive Formalization Methodology:

- **Four-Stage Process**: Strategy representation → linguistic analysis → draft formalizations → final theories
- **Coverage First**: Identifies full breadth before achieving depth/competency
- **Incremental Formalization**: Gradual progression from informal concepts to rigorous logical axioms
- **Contrasts with Traditional Approach**: Avoids early coverage abandonment for competency

3.1.2 2. Strategy Representation Analysis:

- **372 Strategies Collected**: Across 10 planning domains (business, education, warfare, etc.)
- **Analogical Foundation**: Strategies as shared relational structure between analogous planning cases
- **Conceptual Evidence**: Analogies reveal mental representation structures people use
- **Systematic Analysis**: From thousands of words/phrases to 988 unique concepts to 48 representational areas

3.1.3 3. English Language Analysis:

- **Comprehensive Coverage**: Thousands of words/phrases referring to mental states and processes
- **Three-Step Process**: Large-scale brainstorming → linguistic resource expansion → synonymous clustering

- **Finite-State Transducers**: 9,556 states and 85,153 transitions for automatic text annotation
- **PsychoPhrase Game**: Novel methodology for eliciting psychological language references

3.1.4 4. Draft Formalizations:

- **892 Axioms**: Across 30 content theories in first-order logic
- **Productivity Focus**: Liberal use of undefined predicates, relaxed axiomization style
- **Predicate Proliferation**: Multiple forms for single concepts (e.g., 9 forms for "instantiation")
- **Foundation Building**: Created representational basis for commonsense psychology

3.1.5 5. Final Theories Development:

- **Consistency Achievement**: Rigorous semantics and consistent predicate-argument relationships
- **Background Theories**: 16 theories for non-psychological concepts
- **Psychology Theories**: 29 commonsense psychology theories
- **ISO/IEC 24707 Standard**: Common Logic Interchange Format adoption

3.2 Technical Sophistication:

- **Methodological Innovation**: First comprehensive approach to achieve both coverage and competency
- **Scale Achievement**: From informal concepts to formal logical theories across entire domain
- **Linguistic Engineering**: Advanced finite-state transducer technology for concept recognition
- **Logical Formalization**: Systematic progression from informal to formal representations
- **Evaluation Metrics**: 81.5% recall, 95.2% precision in expression identification

3.3 Complexity Distribution:

- **Strategy Collection**: 10 domains, 372 strategies, 8,844 essential feature terms
- **Linguistic Analysis**: 30 representational areas, hundreds of references per concept
- **Draft Theories**: 892 axioms requiring systematic reorganization and refinement
- **Final Implementation**: 16 background + 29 psychology theories with consistent notation

3.4 Conceptual Importance:

This chapter establishes the **methodological foundation** for large-scale knowledge representation in commonsense psychology. The successive formalization approach addresses the fundamental trade-off between coverage and competency that has hindered the field. By prioritizing breadth first, then systematically increasing formalization rigor, the methodology enables comprehensive domain

coverage while achieving logical competency. This represents a paradigm shift from traditional approaches that sacrifice coverage for early competency.

3.5 Cross-Chapter Connections:

- **Chapter 1-2**: Provides methodological foundation for themes introduced in psychology and AI
- **Chapter 4**: Linguistic analysis directly feeds into vocabulary cataloging
- **Parts II-III**: Methodology produces the background and psychology theories formalized later
- **Applied Research**: Finite-state transducers enable corpus analysis applications

3.6 Applications Mentioned:

- **Text Analysis**: Automatic annotation of commonsense psychology concepts in literature
- **Historical Analysis**: Tracking psychological concept evolution (e.g., unconscious desires pre/post-Freud)
- **Developmental Studies**: Children's acquisition of knowledge and belief reasoning
- **Natural Language Processing**: Advanced lexical resource for psychological concept recognition
- **Knowledge Engineering**: Repeatable methodology for other highly interdependent domains

3.7 Notable Design Decisions:

- **Strategy-Based Evidence**: Using planning strategies as surrogate for naturally occurring analogies
- **Linguistic Grounding**: Comprehensive analysis of English psychological vocabulary
- **Productivity Over Rigor**: Initial focus on content generation rather than logical perfection
- **Incremental Refinement**: Four-stage process allowing systematic quality improvement
- **Community Integration**: Publishing intermediate results (finite-state transducers) for broader impact

3.8 Theoretical Significance:

Chapter 3 presents a **revolutionary methodology** for knowledge representation that solves the coverage-competency dilemma plaguing the field. The successive formalization approach demonstrates that comprehensive domain coverage is achievable through systematic progression from informal to formal representations. The strategy-based evidence collection provides novel insights into human mental representation, while the linguistic analysis reveals the remarkable richness of psychological vocabulary. This methodology enables the creation of the comprehensive formal theories presented in subsequent parts of the book.

3.9 Unique Contributions:

3.9.1 Methodological Innovation:

- First systematic approach to achieve both coverage and competency in knowledge representation
- Novel use of planning strategies as evidence for mental representation structures
- Comprehensive linguistic methodology for concept identification and clustering

3.9.2 Empirical Insights:

- Demonstration that 30 of 48 representational areas relate to commonsense psychology
- Evidence for systematic structure in human psychological reasoning across domains
- Quantitative analysis of English psychological vocabulary breadth and organization

3.9.3 Technological Advancement:

- Advanced finite-state transducer technology for automatic psychological concept recognition
- Practical tools enabling large-scale corpus analysis of psychological language
- Repeatable methodology applicable to other complex knowledge domains

3.9.4 Theoretical Foundation:

- Establishes formal basis for comprehensive commonsense psychology theory
- Bridges gap between informal psychological reasoning and formal logical representation
- Provides systematic approach to domain analysis that could transform knowledge engineering

This chapter represents a **methodological breakthrough** that makes the comprehensive formalization of commonsense psychology achievable for the first time, establishing both the theoretical foundation and practical tools necessary for the formal theories developed throughout the remainder of the book.

Chapter 4

Chapter 4: Commonsense Psychology and Language

4.1 Key Features Identified:

4.1.1 1. Comprehensive Vocabulary Catalog:

- **29 Representational Areas**: Complete organization of commonsense psychology concepts
- **Thousands of Expressions**: English words and phrases for mental states and processes
- **Synonymous Groupings**: Conceptually equivalent terms clustered into unified concepts
- **Breadth Over Completeness**: Focus on conceptual distinctions rather than exhaustive listings

4.1.2 2. Knowledge Management (37 concepts):

- **Core Epistemology**: Truth, falsity, belief, knowledge, justification, inference
- **Reasoning Processes**: Adding/removing beliefs, checking inferences, contradiction handling
- **Cognitive Abilities**: Intelligence spectrum, bias patterns, reasoning failures
- **Philosophical Grounding**: Platonic knowledge as justified true belief foundation

4.1.3 3. Memory System (12 concepts):

- **Container Metaphor**: Memory as storage for thoughts with retrieval mechanisms
- **Process Types**: Deliberative vs. passive retrieval, memory cues and reminding
- **Failure Modes**: Retrieval failures, repression, forgetting to remember
- **Temporal Aspects**: Scheduled plan retrieval and failure patterns

4.1.4 4. Envisioning Framework (35 concepts):

- **World Model Manipulation**: Past, present, future, and fictional state reasoning

- **Causal Infrastructure**: Branching timelines, likelihood assessments, consequence propagation
- **Reality Distinctions**: Real-world vs. other-world envisionments
- **Temporal Management**: Advancing world models, validation/invalidation processes

4.1.5 5. Planning and Execution (200+ concepts):

- **Plan Structure**: 20 plan concepts from partial to collaborative to executed plans
- **Goal Management**: 21 goal concepts including themes, conflicts, achievement states
- **Planning Modalities**: Construction, adaptation, design approaches (40+ concepts)
- **Execution Control**: 36 concepts for plan launching, management, and completion

4.1.6 6. Psychological Processes:

- **Other-Agent Reasoning**: Theory of mind, introspection, psychological theories (7 concepts)
- **Decision Making**: Choice enumeration, deliberation, selection, revision (25 concepts)

- **Threat Detection**: Risk assessment, vulnerability recognition, safety evaluation (16 concepts)
- **Emotional States**: 103 emotion categories organized by psychological function

4.2 Technical Sophistication:

- **Systematic Organization**: Hierarchical structure from general categories to specific linguistic expressions
- **Linguistic Precision**: Careful distinction between synonymous and polysemous expressions
- **Conceptual Granularity**: Balance between meaningful distinctions and practical utility
- **Cross-Referenced Structure**: Connections between representational areas clearly indicated
- **Empirical Grounding**: Based on extensive corpus analysis and elicitation studies

4.3 Complexity Distribution:

- **Basic Concepts**: 30% covering fundamental states (knowledge, beliefs, goals)
- **Process Concepts**: 40% covering mental operations (planning, reasoning, deciding)
- **Complex Interactions**: 30% covering multi-agent and temporal reasoning scenarios
- **Emotion Vocabulary**: Hundreds of specific emotional states with low polysemy

4.4 Conceptual Importance:

4.4.1 Foundational Significance:

Establishes the **complete conceptual scope** for formal commonsense psychology theories. The systematic vocabulary provides both the breadth requirements for coverage and the granular

distinctions necessary for competent reasoning. This represents the most comprehensive catalog of psychological language ever assembled for computational purposes.

4.4.2 Theoretical Framework:

Demonstrates that commonsense psychology requires **sophisticated representational infrastructure** spanning knowledge management, temporal reasoning, causal modeling, planning, and emotional processing. The interconnected nature of these concepts reveals why previous attempts at partial formalization failed.

4.4.3 Linguistic Resource:

Creates an unprecedented **lexical foundation** for natural language processing systems requiring psychological understanding. The careful clustering of synonymous expressions enables robust concept recognition across diverse linguistic expressions.

4.5 Cross-Chapter Connections:

- **Chapter 3**: Implements the linguistic analysis methodology described in formalization process
- **Part II (Chapters 5-20)**: Provides target vocabulary for background theory formalization
- **Part III (Chapters 21-49)**: Defines specific concepts formalized in psychology theories
- **Chapter 1-2**: Grounds theoretical motivations in concrete linguistic evidence

4.6 Applications Mentioned:

- **Natural Language Processing**: Lexical resource for interpreting psychological language
- **Knowledge Engineering**: Target specification for formal theory development
- **Computational Linguistics**: Large-scale corpus analysis of psychological concepts
- **Theory Development**: Foundation for alternative formalization approaches
- **Text Analysis**: Automatic identification of mental state references in literature

4.7 Notable Design Decisions:

- **Breadth Prioritization**: Complete coverage over exhaustive synonym listing
- **Functional Organization**: Clustering by psychological process rather than linguistic similarity
- **Granularity Balance**: Meaningful distinctions without excessive subdivision
- **Cross-Part Integration**: Systematic alignment with formal theories in Parts II-III
- **Practical Focus**: Emphasis on computationally useful conceptual distinctions

4.8 Theoretical Significance:

Chapter 4 reveals the **remarkable complexity and systematicity** of commonsense psychological language. The catalog demonstrates that psychological reasoning requires comprehensive infrastructure spanning epistemology, temporal reasoning, planning, and social cognition. The systematic organization reveals deep structural principles organizing human psychological concepts, providing crucial evidence for the formal theories developed in subsequent parts.

4.9 Unique Contributions:

4.9.1 Comprehensive Psychological Vocabulary:

- Most complete catalog of English psychological language ever assembled
- Systematic organization revealing conceptual structure of folk psychology
- Careful distinction between conceptual categories and linguistic expressions

4.9.2 Representational Architecture:

- Identifies 29 fundamental areas required for psychological reasoning
- Demonstrates systematic interconnections between psychological processes
- Provides empirical foundation for formal theory structure

4.9.3 Linguistic Innovation:

- Novel approach to clustering psychological expressions by conceptual function
- Systematic treatment of polysemy and synonymy in psychological language
- Integration of philosophical and empirical approaches to concept identification

4.9.4 Computational Foundation:

- Provides target specification for comprehensive psychological reasoning systems
- Enables systematic evaluation of psychological theory coverage and competency
- Creates foundation for natural language processing of psychological content

4.9.5 Methodological Advancement:

- Demonstrates systematic approach to domain vocabulary identification
- Provides replicable methodology for other complex conceptual domains
- Bridges linguistic analysis and formal knowledge representation

This chapter establishes the **complete conceptual foundation** for formal commonsense psychology, providing both the breadth requirements and granular specifications necessary for the comprehensive logical theories developed throughout the remainder of the book. The systematic vocabulary serves as both theoretical framework and practical resource for computational psychological reasoning.

Chapter 5

Chapter 5: Eventualities and their structure

- **19 axioms total** covering eventualities and their structure
- **3 main sections**: Eventualities and Individuation, Structure of Eventualities, and Generation
- **All background theory** - this is a foundational chapter

5.1 Key Features Identified:

1. **Axiom Schemas**: Several axioms (5.1, 5.3-5.6, 5.11, 5.13) are schemas that need instantiation for each predicate
2. **Reified Predicates**: Heavy use of primed predicates (give', run', go') representing reified eventualities
3. **Core Concepts**:
   - `eventuality` - fundamental concept
   - `argn` - argument structure
   - `gen` - generation relation between eventualities
   - `Rexist` - real existence in the world
4. **Pattern Types**:
   - Argument structure definitions (5.3-5.6)
   - Type constraints (5.7, 5.12, 5.14, 5.16)
   - Recursive definitions (5.10 for arg*)
   - Generation relations (5.2, 5.17-5.21)

5.2 Technical Sophistication:

- **Axiom Schemas**: Multiple schemas (5.1, 5.3-5.6, 5.11, 5.13) requiring instantiation per predicate

- **Reified Predicates**: Extensive use of primed predicates representing eventualities as first-class objects
- **Recursive Definitions**: Complex recursive structure for arg* (5.10) and generation chains
- **Type Integration**: Seamless integration with type theory and existence predicates
- **Foundational Framework**: Provides essential infrastructure for all subsequent theories

5.3 Complexity Distribution:

- Simple: 13 axioms (basic type constraints, simple implications)
- Moderate: 4 axioms (existential definitions, temporal/spatial preservation)
- Complex: 2 axioms (recursive definition of arg*, complete eventuality structure)

5.4 Conceptual Importance:

This chapter provides the **fundamental ontological foundation** for the entire formalization. Eventualities serve as the basic building blocks for representing states, events, and processes in commonsense reasoning. The argument structure framework enables systematic representation of predicate-argument relationships, while the generation relation supports causal and temporal reasoning. Real existence (Rexist) distinguishes between possible and actual eventualities, crucial for modal reasoning.

5.5 Cross-Chapter Connections:

- **Chapter 6 (Sets)**: Eventualities can be organized into sets and collections
- **Chapter 7 (Substitution)**: Substitution operations apply to eventuality arguments
- **Chapter 8 (Logic Reified)**: Logical operations combine and transform eventualities
- **Chapter 15 (Causality)**: Generation relation provides foundation for causal reasoning
- **Chapter 16 (Time)**: Temporal constraints apply to eventuality ordering
- **All Psychology Chapters**: Mental states and processes represented as eventualities

5.6 Applications Mentioned:

- **Natural Language**: Representing verbal predicates and their argument structures
- **Causal Reasoning**: Generation chains for cause-effect relationships
- **Planning**: Actions and goals as structured eventualities
- **Belief Systems**: Mental states as eventualities with argument structure
- **Event Recognition**: Pattern matching against eventuality structures

5.7 Notable Design Decisions:

- **Reification Choice**: Events and states as first-class objects rather than logical formulas
- **Argument Generality**: Universal arg1, arg2, argn structure accommodates all predicates
- **Generation Relation**: Abstract relation supporting multiple forms of derivation
- **Schema Approach**: Template axioms requiring instantiation for specific predicates
- **Existence Integration**: Real existence predicate distinguishes possible from actual

5.8 Theoretical Significance:

Chapter 5 establishes the **core representational framework** that makes comprehensive commonsense reasoning possible. By treating events and states as objects rather than mere logical formulas, the theory enables reasoning about mental attitudes toward states of affairs, causal relationships between events, and temporal ordering of processes. The argument structure provides systematic method for relating entities to the processes they participate in, while generation enables both causal and logical derivation.

5.9 Unique Contributions:

5.9.1 Eventuality Reification:

Treating states and events as objects enables higher-order reasoning about situations, crucial for psychological attitudes like belief, desire, and planning.

5.9.2 Universal Argument Structure:

The arg1, arg2, argn framework provides systematic method for representing predicate-argument relationships across all domains and predicates.

5.9.3 Generation Framework:

Abstract generation relation supports multiple forms of derivation (causal, logical, temporal) essential for commonsense reasoning.

5.9.4 Ontological Foundation:

Provides the basic building blocks upon which all subsequent theories—both background and psychological—can be constructed systematically.

This chapter represents the **foundational architectural decision** that enables the comprehensive approach to commonsense psychology, providing the representational infrastructure necessary for all subsequent formal theories.

Chapter 6

Chapter 6: Tradicional set theory

- **22 axioms total** covering traditional set theory operations
- **Single section** but comprehensive coverage of set theory fundamentals
- **All background theory** - foundational mathematical concepts

6.1 Key Features Identified:

1. **Set Theory Foundation**: Standard Zermelo-Fraenkel set theory without Axiom of Infinity
2. **Equality Properties**: Axioms 6.3-6.5 establish reflexivity, symmetry, and transitivity of equality
3. **Set Operations**: Complete coverage of:
   - Basic operations: union, intersection, difference
   - Construction: addElt, deleteElt, replaceElt
   - Special sets: null, singleton, doubleton, power set
   - Relations: subset, proper subset, disjoint
4. **Pattern Types**:
   - Set construction and manipulation (6.8, 6.11, 6.12)
   - Standard set operations (6.13, 6.15, 6.16)
   - Recursive definitions (6.21 for cardinality)
   - Type constraints and definitions
5. **Complexity Distribution**:
   - Simple: 6 axioms (basic properties, simple definitions)
   - Moderate: 8 axioms (medium complexity definitions)
   - Complex: 8 axioms (multi-quantifier definitions like equality, union, intersection)

6.2 Technical Sophistication:

- **Mathematical Rigor**: Full Zermelo-Fraenkel set theory implementation with careful attention to foundational principles
- **Operational Completeness**: Comprehensive coverage of standard set operations (union, intersection, difference, power set)
- **Construction Methods**: Systematic element manipulation operations (addElt, deleteElt, replaceElt)
- **Type Integration**: Typed sets connecting set theory to eventuality framework from Chapter 5
- **Recursive Definitions**: Complex recursive structure for cardinality calculations

6.3 Complexity Distribution:

- Simple: 6 axioms (basic properties like reflexivity, symmetry of equality)
- Moderate: 8 axioms (medium complexity definitions for operations like union, intersection)
- Complex: 8 axioms (multi-quantifier definitions, particularly set equality and comprehensive operations)

6.4 Conceptual Importance:

Provides the **essential mathematical foundation** for organizing and manipulating collections of objects throughout the commonsense reasoning system. Set theory enables systematic representation of groups, categories, and relationships between entities. The mathematical rigor ensures consistent reasoning about collections, while the integration with eventualities connects abstract mathematical operations to concrete reasoning scenarios.

6.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Typed sets organize eventualities into meaningful collections
- **Chapter 7 (Substitution)**: Set operations combine with substitution for systematic replacements
- **Chapter 8 (Logic)**: Logical operations apply to sets of eventualities and propositions
- **Chapter 9 (Composite Entities)**: Sets provide foundation for complex entity construction
- **Planning Chapters**: Goal sets and plan collections use set-theoretic operations
- **Psychology Chapters**: Belief sets, knowledge collections organized using set theory

6.6 Applications Mentioned:

- **Knowledge Organization**: Grouping related concepts and beliefs into coherent collections

- **Planning Systems**: Managing sets of goals, actions, and constraints
- **Categorization**: Systematic representation of conceptual hierarchies and taxonomies
- **Resource Management**: Organizing available tools, agents, and capabilities
- **Logical Reasoning**: Set-based operations for inference and contradiction detection

6.7 Notable Design Decisions:

- **ZF Foundation**: Zermelo-Fraenkel axioms without Axiom of Infinity for computational tractability
- **Operational Focus**: Emphasis on constructive operations rather than purely abstract properties
- **Type Integration**: Connecting mathematical set theory to domain-specific eventuality types
- **Recursive Approach**: Enabling complex cardinality calculations through recursive definitions
- **Practical Operations**: Including element manipulation for dynamic set modification

6.8 Theoretical Significance:

Chapter 6 establishes the **mathematical infrastructure** necessary for systematic reasoning about collections and categories. Set theory provides the formal foundation for organizing knowledge, managing complexity, and ensuring consistent reasoning about groups of related objects. The integration with the eventuality framework creates a unified system where mathematical rigor supports practical commonsense reasoning applications.

6.9 Unique Contributions:

6.9.1 Foundational Integration:

Seamless connection between abstract mathematical set theory and concrete eventuality-based reasoning, enabling mathematical operations on domain-specific objects.

6.9.2 Operational Completeness:

Comprehensive coverage of both theoretical set properties and practical manipulation operations needed for dynamic reasoning systems.

6.9.3 Type-Theoretic Bridge:

Novel integration of typed sets with eventuality framework, creating systematic method for organizing domain knowledge.

6.9.4 Computational Tractability:

Careful selection of ZF axioms ensuring mathematical rigor while maintaining computational feasibility for reasoning systems.

This chapter provides the **mathematical backbone** that enables systematic organization and manipulation of knowledge throughout the commonsense reasoning system, establishing essential infrastructure for all subsequent theories requiring collection-based reasoning.

Chapter 7

Chapter 7: Substitution, Typical elements and Instances

- **19 axioms total** covering substitution, typical elements, functional dependencies, and instantiation
- **5 main sections**: Substitution, Typical Elements, Handling Thorny Issues, Functional Dependencies, and Instances
- **All background theory** - advanced foundational concepts

7.1 Key Features Identified:

1. **Substitution Framework**:
   - Axiom 7.1 is extremely complex, defining recursive substitution between eventualities
   - Axiom 7.2 defines composite substitutions (subst2)
2. **Typical Elements**:
   - Reification of universally quantified variables as "typical elements"
   - Property inheritance from typical elements to real set members (7.4)
   - Defined sets (dset) with their typical elements (7.5-7.6)
3. **Technical Safeguards**:
   - Axioms 7.7-7.8 prevent problematic applications involving set membership and typical elements
   - These "blocking" predicates solve Russell paradox-type issues
4. **Functional Dependencies**:
   - Captures existentially quantified variables through functional dependencies (fd)
   - Skolem functions (skfct) and their ranges (rangeFd)
   - Dependency inheritance for partial instantiations (7.13)
5. **Instantiation Framework**:

- Parameters of abstract eventualities (7.16)
- Partial and complete instantiation (7.17-7.18)
- Type instantiation with holdsFor (7.19)

6. **Complexity Distribution**:
   - Simple: 3 axioms (basic definitions)
   - Moderate: 13 axioms (most of the functional machinery)
   - Complex: 3 axioms (recursive substitution, dependency inheritance, partial instantiation)

7.2 Technical Sophistication:

- **Recursive Substitution**: Axiom 7.1 implements complex recursive substitution with nested quantifiers and multiple argument handling
- **Typical Element Reification**: Novel approach to handling universally quantified variables as first-class objects
- **Paradox Resolution**: Technical safeguards (blocking predicates) preventing Russell paradox-type issues in set membership
- **Functional Dependency Framework**: Systematic handling of existentially quantified variables through Skolem functions
- **Instantiation Hierarchy**: Complete framework for managing abstract to concrete object transitions

7.3 Complexity Distribution:

- Simple: 3 axioms (basic definitions for ranges, parameters, type instantiation)
- Moderate: 13 axioms (functional dependencies, typical element properties, instantiation machinery)
- Complex: 3 axioms (recursive substitution 7.1, dependency inheritance 7.13, partial instantiation 7.17)

7.4 Conceptual Importance:

This chapter provides **critical infrastructure** for reasoning about generality and specificity in commonsense knowledge. The substitution framework enables systematic replacement of abstract concepts with concrete instances, essential for applying general knowledge to specific situations. Typical elements bridge the gap between universal statements and particular cases, while functional dependencies handle existential relationships. This machinery is fundamental for practical reasoning systems that must apply general principles to specific scenarios.

7.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Substitution operates on eventuality arguments and structures
- **Chapter 6 (Sets)**: Typical elements provide systematic method for reasoning about set properties
- **Chapter 8 (Logic)**: Logical operations combine with substitution for complex inferences
- **Chapter 9 (Composite Entities)**: Instantiation framework supports part-whole reasoning
- **Planning Chapters**: Abstract plans instantiated to specific execution contexts
- **Psychology Chapters**: General psychological principles applied to specific mental states

7.6 Applications Mentioned:

- **Knowledge Application**: Converting general principles to specific situation reasoning
- **Plan Instantiation**: Transforming abstract strategies into executable action sequences
- **Concept Specification**: Moving from general categories to specific instances
- **Variable Binding**: Systematic handling of quantified variables in logical reasoning
- **Pattern Matching**: Matching abstract patterns against concrete situations

7.7 Notable Design Decisions:

- **Reified Variables**: Treating typical elements as objects rather than logical variables
- **Recursive Framework**: Enabling complex nested substitutions for deep structural changes
- **Paradox Prevention**: Including blocking predicates to avoid set-theoretic paradoxes
- **Functional Approach**: Using Skolem functions for systematic existential handling
- **Partial Instantiation**: Supporting gradual specification of abstract concepts

7.8 Theoretical Significance:

Chapter 7 addresses the **fundamental challenge** of connecting abstract knowledge to concrete application. The substitution and instantiation frameworks enable reasoning systems to apply general principles systematically to specific cases, essential for practical intelligence. The typical element approach provides novel solution to universal quantification in reified logic, while functional dependencies handle existential relationships systematically. This infrastructure is crucial for the psychological theories that follow, which must relate general principles to specific mental states and processes.

7.9 Unique Contributions:

7.9.1 Reified Substitution Framework:

Revolutionary approach to substitution in reified logic, enabling systematic replacement operations on first-class eventuality objects rather than logical formulas.

7.9.2 Typical Element Innovation:

Novel method for handling universal quantification by reifying typical elements, bridging gap between logical variables and concrete objects.

7.9.3 Paradox-Safe Design:

Careful technical safeguards preventing Russell paradox issues while maintaining expressive power for practical reasoning applications.

7.9.4 Functional Dependency System:

Systematic framework for handling existential quantification through Skolem functions, enabling practical reasoning about unknown entities.

7.9.5 Instantiation Architecture:

Comprehensive framework for managing transitions from abstract to concrete representations, essential for applying general knowledge to specific situations.

This chapter establishes the **essential bridge** between abstract knowledge representation and concrete application, providing the technical infrastructure necessary for practical commonsense reasoning systems that must systematically apply general principles to specific real-world scenarios.

Chapter 8

Chapter 8: Logic reified

- **13 axioms total** covering reified logical operators
- **3 main sections**: Conjunction, Negation, and Disjunction and Implication
- **All background theory** - foundational logical machinery

8.1 Key Features Identified:

1. **Reified Logic**:
   - Standard logical operators (and, or, not, imply) are treated as first-class predicates
   - Primed versions (and', or', not', imply') are the basic reified forms
   - This allows logical relationships to be objects of thought and belief
2. **Modal Existence**:
   - Heavy use of `Rexist` predicate to handle when logical combinations actually hold
   - Axiom 8.1: Conjunction exists when both conjuncts exist
   - Axiom 8.3: Negation exists when the negated thing doesn't exist
   - Axiom 8.6: Disjunction exists when at least one disjunct exists
3. **Type vs Token Distinction**:
   - Axiom 8.4 introduces `nott'` for negating eventuality types rather than tokens
   - Important for handling contextual negation in natural language
4. **Set Extensions**:
   - Axiom 8.2: Sets exist when all members exist (conjunction interpretation)
   - Axiom 8.7: Disjunction extended to sets of eventualities
   - Axiom 8.8: Implications can take sets as antecedents
5. **Inference Rules**:
   - Axiom 8.10: Modus ponens for reified implications
   - Axiom 8.11: Transitivity of implication
   - Axiom 8.12: Definition of inconsistent sets
   - Axiom 8.13: Minimal proof relationship

6. **Complexity Distribution**:
   - Simple: 3 axioms (basic inference rules, type constraints)
   - Moderate: 10 axioms (most of the logical definitions)
   - Complex: 0 axioms (relatively straightforward chapter)

8.2 Technical Sophistication:

- **Reified Logic Framework**: Standard logical operators (and, or, not, imply) treated as first-class predicates
- **Modal Existence Handling**: Sophisticated use of Rexist predicate for determining when logical combinations hold
- **Type/Token Distinction**: Advanced treatment of nott' for negating eventuality types vs. tokens
- **Set Extensions**: Systematic extension of logical operations to sets of eventualities
- **Inference Integration**: Complete modus ponens and transitivity rules for reified implications

8.3 Complexity Distribution:

- Simple: 3 axioms (basic inference rules, type constraints)
- Moderate: 10 axioms (logical definitions, set extensions, modal existence conditions)
- Complex: 0 axioms (relatively straightforward logical machinery)

8.4 Conceptual Importance:

This chapter provides **essential logical infrastructure** for reasoning about reasoning itself. By reifying logical operators, the system can treat logical relationships as objects of thought and belief, crucial for psychological reasoning about mental states. The modal existence framework enables systematic handling of when logical combinations actually hold in the world, while the type/token distinction supports contextual interpretation of logical statements. This foundation enables agents to reason about their own and others' logical reasoning processes.

8.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Logical operations apply to reified eventualities and their combinations
- **Chapter 6 (Sets)**: Set-based extensions of logical operations for collective reasoning
- **Chapter 7 (Substitution)**: Logical operations interact with substitution for complex inferences
- **Knowledge Chapters**: Belief and knowledge states represented through reified logical relationships

- **Planning Chapters**: Logical reasoning about plan conditions, effects, and constraints
- **Psychology Chapters**: Mental reasoning processes modeled through reified logical operations

8.6 Applications Mentioned:

- **Meta-Reasoning**: Reasoning about logical relationships as objects of thought
- **Belief Systems**: Representing beliefs about logical relationships and their validity
- **Plan Reasoning**: Logical analysis of plan preconditions, effects, and consistency
- **Contradiction Detection**: Systematic identification of inconsistent belief sets
- **Proof Construction**: Building and analyzing logical arguments as structured objects

8.7 Notable Design Decisions:

- **Reification Choice**: Logical operators as predicates rather than meta-logical constructs
- **Existence Integration**: Modal existence conditions determining when logical combinations hold
- **Type Sensitivity**: Separate negation operators for types vs. tokens supporting contextual reasoning
- **Set Generalization**: Extending logical operations to collections of eventualities
- **Inference Preservation**: Maintaining standard logical inference rules in reified form

8.8 Theoretical Significance:

Chapter 8 establishes the **logical reasoning infrastructure** necessary for sophisticated commonsense psychology. By reifying logical operators, agents can reason about logical relationships as mental objects, essential for modeling beliefs, knowledge, and reasoning processes. The systematic integration with modal existence enables realistic reasoning about when logical relationships actually hold, while the type/token distinction supports context-sensitive logical interpretation essential for natural language understanding.

8.9 Unique Contributions:

8.9.1 Logic Reification Framework:

Systematic treatment of logical operators as first-class objects, enabling meta-reasoning about logical relationships and their properties.

8.9.2 Modal Logic Integration:

Novel combination of reified logic with modal existence, providing realistic framework for reasoning about when logical relationships hold.

8.9.3 Type/Token Logical Distinction:

Innovative handling of contextual negation through separate operators for eventuality types versus specific tokens.

8.9.4 Set-Theoretic Extensions:

Comprehensive generalization of logical operations to sets of eventualities, supporting collective reasoning and complex inference patterns.

8.9.5 Inference Rule Preservation:

Careful maintenance of standard logical inference rules (modus ponens, transitivity) within reified framework, ensuring logical soundness.

This chapter provides the **logical foundation** that bridges formal reasoning and psychological modeling, enabling agents to reason systematically about their own and others' logical reasoning processes while maintaining the rigor necessary for sound inference.

Chapter 9

Chapter 9: Functions and sequences

- **19 axioms total** covering pairs, functions, and sequences
- **2 main sections**: Pairs and Functions, and Sequences
- **All background theory** - mathematical foundations

9.1 Key Features Identified:

1. **Ordered Pairs**:
   - Axiom 9.1: Definition of pairs through existential quantification
   - Axiom 9.2: Structural properties (first, second elements)
   - Axioms 9.3-9.4: Equality and uniqueness properties
2. **Functions**:
   - Axiom 9.5: Complex definition of function0 as sets of pairs with constraints
   - Functions map from domain to range with unique values
   - Axioms 9.6-9.7: Domain and range definitions
   - Axiom 9.8: Function mapping (f(x) = y)
   - Axiom 9.9: Connection to Skolem functions from Chapter 7
3. **Sequences**:
   - Axiom 9.10: Integer intervals (sets of consecutive integers)
   - Axiom 9.11: Sequences as functions from {1,...,n} to any range
   - Axiom 9.12: Length definition
   - Axioms 9.13-9.14: nth and first element access
   - Axiom 9.15: Rest operation (sequence after removing first element)
   - Axioms 9.16-9.19: Last element, membership, ordering, and succession
4. **Mathematical Dependencies**:
   - References to positive integers, successor, leq, lt from Chapter 13
   - Builds on set theory from Chapter 6
   - Connects to Skolem functions from Chapter 7

5. **Complexity Distribution**:
   - Simple: 8 axioms (basic definitions and properties)
   - Moderate: 9 axioms (medium complexity definitions)
   - Complex: 2 axioms (function definition 9.5, sequence rest 9.15)
6. **Notable Features**:
   - Axiom 9.5 is particularly complex with multiple nested quantifiers
   - Overloaded predicates: `first` works for both pairs and sequences
   - Rich sequence operations for temporal/causal reasoning later

9.2 Technical Sophistication:

- **Set-Theoretic Foundations**: Functions rigorously defined as constrained sets of ordered pairs with unique value mappings
- **Ordered Pair Framework**: Complete implementation with existential quantification and structural properties
- **Sequence Operations**: Comprehensive suite of operations (nth, first, rest, last, succession) for temporal reasoning
- **Mathematical Integration**: Systematic connection to arithmetic (Chapter 13) and Skolem functions (Chapter 7)
- **Overloaded Predicates**: Elegant reuse of predicates like `first` across pairs and sequences

9.3 Complexity Distribution:

- Simple: 8 axioms (basic definitions, element access, length properties)
- Moderate: 9 axioms (domain/range definitions, membership, ordering operations)
- Complex: 2 axioms (function definition 9.5 with multiple nested quantifiers, sequence rest operation 9.15)

9.4 Conceptual Importance:

This chapter provides **essential mathematical infrastructure** for representing structured relationships and temporal sequences in commonsense reasoning. Functions enable systematic representation of mappings between entities, essential for causal relationships, state transformations, and cognitive processes. Sequences support temporal ordering, narrative structure, and procedural reasoning. The set-theoretic foundation ensures mathematical rigor while the comprehensive operation suite enables practical reasoning applications.

9.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Functions map between eventuality arguments and their relationships
- **Chapter 6 (Sets)**: Functions built on set-theoretic foundations with ordered pair structures
- **Chapter 7 (Substitution)**: Skolem functions connect logical dependencies to mathematical functions
- **Chapter 13 (Arithmetic)**: Integer sequences and mathematical operations integration
- **Chapter 15 (Causality)**: Sequences represent causal chains and temporal ordering
- **Chapter 16 (Time)**: Temporal sequences for representing event ordering and duration

9.6 Applications Mentioned:

- **Temporal Reasoning**: Sequences representing event orderings and narrative structures
- **Causal Modeling**: Function mappings for cause-effect relationships and state transitions
- **Procedural Representation**: Sequences encoding step-by-step processes and plans
- **Data Structures**: Mathematical foundation for complex structured representations
- **Parameter Binding**: Functions mapping abstract concepts to specific values

9.7 Notable Design Decisions:

- **Set-Theoretic Definition**: Functions as constrained sets of pairs ensuring mathematical rigor
- **Predicate Overloading**: Systematic reuse of predicates across different mathematical structures
- **Comprehensive Operations**: Complete suite of sequence operations for practical reasoning needs
- **Integer Integration**: Forward references to arithmetic enabling mathematical consistency
- **Skolem Connection**: Bridging logical dependencies and mathematical function concepts

9.8 Theoretical Significance:

Chapter 9 establishes the **mathematical foundation** necessary for representing structured relationships and temporal processes in commonsense reasoning. Functions provide the formal basis for representing mappings, transformations, and dependencies that pervade psychological and causal reasoning. Sequences enable systematic representation of temporal order, procedural knowledge, and narrative structure essential for practical intelligence. The rigorous mathematical foundation ensures consistency while supporting the complex structured reasoning required for psychology and planning.

9.9 Unique Contributions:

9.9.1 Unified Mathematical Framework:

Systematic integration of functions and sequences providing complete mathematical infrastructure for structured commonsense reasoning.

9.9.2 Set-Theoretic Rigor:

Rigorous mathematical definitions ensuring consistency while maintaining practical utility for complex reasoning applications.

9.9.3 Sequence Operation Suite:

Comprehensive collection of sequence operations optimized for temporal and procedural reasoning in commonsense domains.

9.9.4 Cross-Domain Integration:

Elegant connections between logical dependencies (Skolem functions) and mathematical structures, bridging formal logic and mathematics.

9.9.5 Practical Mathematical Design:

Mathematical foundations designed specifically for commonsense reasoning applications rather than pure mathematical abstraction.

This chapter provides the **mathematical backbone** essential for representing the structured relationships, temporal sequences, and functional dependencies that underlie sophisticated commonsense reasoning about processes, causation, and psychological states.

Chapter 10

Chapter 10: Composite entities

22 axioms total covering composite entities, figure-ground relations, and patterns 4 main sections: Definitions, Simple Examples, Figure-Ground Relation, and Patterns and Their Instances All background theory - foundational concepts for complex structures

10.1 Key Features Identified:

1. Composite Entity Framework:

Axiom 10.1: Core definition - entities with components, properties, and relations Axioms 10.2-10.3: Components must be non-empty sets Axiom 10.4: Aggregates as simple two-component entities Axioms 10.5-10.6: componentOrWhole and externalTo relationships

2. Properties and Relations Constraints:

Axiom 10.7: onlyarg* - recursive definition for single-argument properties Axiom 10.8: Properties must have onlyarg* that's a component or whole Axiom 10.9: Relations must involve a component/whole and something else Axioms 10.10-10.11: Single relation and combined property/relation predicates

3. Examples as Composite Entities:

Axiom 10.12: Sets as composite entities (members as components) Axiom 10.13: Pairs as composite entities (first/second elements, relations) Axiom 10.14: Sequences as composite entities (elements + ordering relations)

4. Figure-Ground Relation:

Axiom 10.15: Basic constraints on the 'at' relation Axioms 10.16-10.17: Two equivalent definitions of 'ground' (shared properties) Axiom 10.18: 'at' relation requires ground as third argument

5. Pattern System:

Axiom 10.19: Patterns contain type elements as components Axiom 10.20: Pattern parameters are the type element components Axiom 10.21: Pattern instances replace all type elements with real entities Axiom 10.22: Incomplete instances have some but not all parameters instantiated

6. Complexity Distribution:

Simple: 6 axioms (basic definitions, constraints) Moderate: 10 axioms (medium complexity definitions with quantifiers) Complex: 6 axioms (nested quantifiers, recursive definitions, pattern instances)

10.2 Notable Technical Features:

Recursive Definitions: onlyarg* (10.7) recurses through eventuality arguments Dual Ground Definitions: Axioms 10.16-10.17 provide equivalent formulations using substitution vs. typical elements Complex Pattern Logic: Axioms 10.21-10.22 handle complete and partial instantiation with property/relation preservation Reified Examples: Axioms 10.12-10.14 use reified predicates (set', pair0', sequence')

10.3 Conceptual Importance:

Figure-Ground: Fundamental cognitive relationship from spatial reasoning Composite Structure: Foundation for understanding complex objects, events, and information structures Pattern Templates: Enables reasoning about types and their instances Uniform Treatment: Physical objects, events, and abstract structures all treated uniformly

10.4 Cross-Chapter Connections:

Builds on sets (Chapter 6), substitution and typical elements (Chapter 7) Uses reified predicates from eventualities (Chapter 5) Figure-ground relation will be crucial for scales and spatial reasoning Pattern system connects to functional dependencies and instantiation

10.5 Domain Applications:

The chapter mentions diverse applications:

Physical: doors, cups, telephones, chairs, automobiles Biological: trees, bees, persons Events: hikes, erosion, concerts Information: equations, sentences, theories, schedules Mixed: books (physical + informational)

10.6 Technical Sophistication:

- **Composite Entity Framework**: Systematic treatment of complex structures with components, properties, and relations

- **Recursive Definitions**: onlyarg* recursively processes eventuality arguments for property constraints
- **Dual Ground Definitions**: Two equivalent formulations using substitution vs. typical elements
- **Pattern Template System**: Complete framework for type instantiation with partial and complete instances
- **Reified Structure Integration**: Seamless connection with reified predicates from eventuality framework

10.7 Complexity Distribution:

- Simple: 6 axioms (basic definitions, constraints on components and relationships)
- Moderate: 10 axioms (medium complexity with quantifiers, figure-ground relations)
- Complex: 6 axioms (recursive definitions, pattern instances, complete instantiation logic)

10.8 Conceptual Importance:

This chapter provides the **architectural foundation** for representing complex structured entities in commonsense reasoning. Composite entities enable systematic representation of physical objects, events, and abstract structures through uniform component-property-relation framework. The figure-ground relation captures fundamental spatial and attentional relationships, while the pattern system enables reasoning about types, templates, and their instantiations. This infrastructure is essential for representing the rich structured world of commonsense psychology.

10.9 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Uses reified predicates for structural representation
- **Chapter 6 (Sets)**: Sets, pairs, and sequences as examples of composite entities
- **Chapter 7 (Substitution)**: Pattern instantiation uses substitution and typical elements
- **Chapter 12 (Scales)**: Figure-ground relation crucial for spatial and measurement scales
- **Chapter 18 (Space)**: Figure-ground provides foundation for spatial reasoning
- **Psychology Chapters**: Mental structures and processes represented as composite entities

10.10 Applications Mentioned:

- **Physical Objects**: doors, cups, telephones, chairs, automobiles with component structure
- **Biological Entities**: trees, bees, persons as structured living systems
- **Event Structures**: hikes, erosion, concerts as temporally structured processes
- **Information Structures**: equations, sentences, theories, schedules as abstract composites
- **Mixed Entities**: books combining physical and informational aspects

- **Pattern Recognition**: Template matching and type-instance reasoning

10.11 Notable Design Decisions:

- **Uniform Treatment**: Physical, biological, event, and abstract structures handled identically
- **Component Constraints**: Non-empty component sets ensuring meaningful structure
- **Property Restrictions**: Properties must apply to components or wholes, preventing ill-formed structures
- **Dual Ground Formulation**: Alternative equivalent definitions supporting different reasoning styles
- **Pattern Graduality**: Supporting both complete and partial instantiation for flexible reasoning

10.12 Theoretical Significance:

Chapter 10 establishes the **structural foundation** for representing complex entities in commonsense reasoning. The composite entity framework provides systematic method for decomposing and reasoning about complex structures, essential for understanding physical objects, events, and abstract concepts. The figure-ground relation captures fundamental perceptual and attentional distinctions, while the pattern system enables sophisticated type-instance reasoning. This infrastructure is crucial for psychological theories involving structured mental representations and complex reasoning processes.

10.13 Unique Contributions:

10.13.1 Unified Structural Framework:

Systematic approach to representing diverse complex entities (physical, biological, abstract, temporal) through uniform component-property-relation structure.

10.13.2 Figure-Ground Integration:

Formal treatment of fundamental cognitive distinction between figure and ground, essential for spatial and attentional reasoning.

10.13.3 Pattern Template System:

Comprehensive framework for type-instance reasoning with support for partial instantiation and gradual specification.

10.13.4 Recursive Property Constraints:

Sophisticated onlyarg* mechanism ensuring properties apply appropriately to components and wholes.

10.13.5 Cross-Domain Applicability:

Unified treatment enabling consistent reasoning across physical objects, events, abstract structures, and mixed entities.

This chapter provides the **essential structural infrastructure** that enables sophisticated reasoning about the complex, hierarchically organized entities that populate commonsense knowledge and psychological reasoning, establishing the foundation for systematic representation of structured mental content.

Chapter 11

Chapter 11: Defeasibility

- **14 axioms total** covering defeasible reasoning and nonmonotonic logic
- **Single main section** but introduces crucial concepts for the entire book
- **13 examples, 1 psychology** - mostly illustrative examples with one psychological axiom

11.1 Key Features Identified:

1. **Classical vs. Defeasible Reasoning**:
   - Axiom 11.1: Classical "birds fly" rule that leads to contradictions
   - Axiom 11.2: Defeasible version using etc1 predicate
   - Shows the problem with monotonic reasoning in commonsense knowledge
2. **Extended Bird Example**:
   - Axioms 11.3-11.4: Emus are birds but don't fly (the exception case)
   - Axioms 11.7-11.8: Additional defeasible properties (two legs, feathers)
   - Shows how multiple defeasible rules can coexist
3. **Alternative Approaches**:
   - Axioms 11.5-11.6: Circumscriptive logic using "abnormal" (ab) predicates
   - Different way to encode the same defeasible reasoning
4. **Etc Predicate System**:
   - Different etc predicates for different axioms (etc1, etc2, etc3, etc4)
   - Axiom 11.9: Biconditional relationship between etc3 and feathered
   - Axiom 11.10: Relationships between etc predicates
   - Axiom 11.12: Psychology example - mothers love children
5. **Notation Development**:
   - Axiom 11.13: Full form with indexed etc predicate (etc.11.14)
   - Axiom 11.14: Abbreviated form using simple "etc"
   - The abbreviation is for readability - must be expanded for automated reasoning
6. **Technical Insights**:

- **Nonmonotonic Logic**: New information can defeat previous conclusions
- **Weighted Abduction**: Framework for handling costs of assumptions
- **Indexing**: Each axiom gets its own etc predicate to avoid interference
- **Variable Scope**: etc predicates must include all universally quantified variables

7. **Complexity Distribution**:
   - Simple: 11 axioms (most are straightforward defeasible rules)
   - Moderate: 3 axioms (biconditional, general patterns)
   - Complex: 0 axioms (conceptually sophisticated but formally simple)

11.2 Conceptual Importance:

This chapter is foundational for the entire book because: - Most commonsense knowledge is defeasible, not absolute - The "etc" notation appears throughout the psychology theories - Enables realistic modeling of human reasoning under uncertainty - Bridges formal logic and practical knowledge representation

11.3 Technical Sophistication:

- **Nonmonotonic Logic Framework**: Complete system for defeasible reasoning where new information can defeat previous conclusions
- **Indexed Etc Predicates**: Sophisticated indexing system (etc1, etc2, etc.11.14) preventing axiom interference
- **Variable Scope Management**: Etc predicates must include all universally quantified variables for proper operation
- **Alternative Encodings**: Both etc-based and circumscriptive logic approaches supported
- **Weighted Abduction Integration**: Framework for handling assumption costs and plausibility

11.4 Complexity Distribution:

- Simple: 11 axioms (straightforward defeasible rules with etc predicates)
- Moderate: 3 axioms (biconditional relationships, general patterns between etc predicates)
- Complex: 0 axioms (conceptually sophisticated but formally straightforward)

11.5 Conceptual Importance:

This chapter provides the **fundamental reasoning infrastructure** for realistic commonsense knowledge representation. Since most commonsense knowledge is defeasible rather than absolute, the etc predicate system enables agents to make reasonable assumptions while remaining open to contradictory evidence. The framework bridges formal logic and practical reasoning, enabling systems

to function with incomplete information while maintaining logical coherence when assumptions are defeated.

11.6 Cross-Chapter Connections:

- **Chapters 21-49 (Psychology)**: Extensive use of etc predicates throughout all psychological theories
- **Chapter 15 (Causality)**: Defeasible causal reasoning using etc framework
- **Chapter 16 (Time)**: Temporal reasoning with defeasible assumptions
- **All Background Chapters**: Defeasible extensions of mathematical and logical foundations
- **Planning Chapters**: Defeasible plan reasoning and assumption management

11.7 Applications Mentioned:

- **Commonsense Reasoning**: "Birds fly" with exceptions like emus, penguins
- **Biological Classification**: Defeasible properties (feathers, two legs) with exceptions
- **Psychological Reasoning**: "Mothers love their children" with rare exceptions
- **Natural Language Understanding**: Default interpretations with contextual overrides
- **Planning Systems**: Defeasible assumptions about action outcomes and preconditions

11.8 Notable Design Decisions:

- **Indexed Predicates**: Separate etc predicates for each axiom preventing logical interference
- **Abbreviation System**: Readable "etc" notation expandable to formal indexed predicates
- **Alternative Approaches**: Support for both etc-based and abnormality-based circumscription
- **Variable Inclusion**: Etc predicates must scope over all relevant universally quantified variables
- **Biconditional Relationships**: Systematic connections between related etc predicates

11.9 Theoretical Significance:

Chapter 11 establishes the **nonmonotonic foundation** essential for practical commonsense reasoning systems. The defeasible reasoning framework enables agents to make reasonable default assumptions while maintaining the ability to revise conclusions when contradictory evidence emerges. This capability is fundamental to human-like reasoning and essential for the psychological theories that follow, which must handle the inherent uncertainty and context-sensitivity of mental processes.

11.10 Unique Contributions:

11.10.1 Comprehensive Defeasible Framework:

Complete system for nonmonotonic reasoning specifically designed for commonsense knowledge representation with practical indexing and scoping mechanisms.

11.10.2 Etc Predicate Innovation:

Novel indexing system preventing axiom interference while maintaining readability through abbreviation conventions.

11.10.3 Alternative Logic Integration:

Support for multiple nonmonotonic approaches (etc-based and circumscriptive) enabling flexible reasoning strategies.

11.10.4 Psychological Applicability:

Framework specifically designed to support the defeasible nature of psychological reasoning and commonsense knowledge.

11.10.5 Weighted Abduction Connection:

Integration with broader reasoning framework for handling assumption costs and plausibility in practical systems.

This chapter provides the **essential nonmonotonic foundation** that makes realistic commonsense reasoning possible, enabling the sophisticated defeasible reasoning about psychology and human behavior developed throughout the remainder of the book.

Chapter 12

Chapter 12: Scales

- **38 axioms total** covering scales, partial orderings, and qualitative structure
- **4 main sections**: Basics, Scale-to-Scale Functions, Constructing Scales, and Qualitative Structure on Scales
- **All background theory** - mathematical and conceptual foundations

12.1 Key Features Identified:

1. **Scale Foundation**:
   - Axiom 12.1: Complex definition of scales as composite entities with partial orderings
   - Axioms 12.2-12.3: Basic constraints and abbreviations (inScale)
   - Axioms 12.5-12.7: Standard partial ordering properties (antireflexive, antisymmetric, transitive)
2. **Scale Relations**:
   - Axiom 12.4: Definition of lts (less than scale) using substitution
   - Axioms 12.8-12.14: Related ordering relations (gts, leqs, geqs)
   - Axioms 12.9-12.10: Arguments can be points on scale or entities at points
   - Axiom 12.11: onScale covers both cases
3. **Scale Structure**:
   - Axioms 12.15-12.16: Top and bottom elements
   - Axiom 12.17: Subscales with restricted components and ordering
   - Axiom 12.18: Reverse scales
   - Axiom 12.20: Total orderings (all elements comparable)
   - Axiom 12.21: Scales are grounds for the at relation
4. **Scale Functions**:
   - Axiom 12.22: Extended function definition for scales
   - Axiom 12.23: Monotone-increasing functions preserve ordering
   - Axiom 12.24: Functions “into” vs “onto”

5. **Scale Construction**:
   - Axiom 12.25: scaleDefinedBy for building scales
   - Axioms 12.26-12.27: Examples - sequences and sets of sets as scales
   - Axiom 12.28: Subset-consistent scales
   - Axiom 12.30: Composite scales from multiple orderings
6. **Qualitative Structure**:
   - Axioms 12.31-12.33: Hi, Md, Lo regions as subscales
   - Axioms 12.34-12.36: Regional extrema and ordering relationships
   - Axiom 12.37: scaleFor connects properties to scales ("tall" → height scale)
   - Axiom 12.38: Hi region properties are defeasibly goal-relevant
7. **Complexity Distribution**:
   - Simple: 14 axioms (basic properties, regional constraints)
   - Moderate: 17 axioms (definitions with multiple conditions)
   - Complex: 7 axioms (scale definition, ordering preservation, scale construction)

12.2 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Qualitative Reasoning**: Hi/Md/Lo regions enable approximate judgments - **Natural Language**: Connects comparative ("taller") and absolute ("tall") adjectives - **Functionality**: Links scalar properties to goal achievement (Axiom 12.38) - **Mathematical Foundation**: Rigorous treatment of partial orderings and scale operations

12.3 Cross-Chapter Connections:

- Builds on composite entities (Chapter 10) - scales are composite entities
- Uses substitution framework (Chapter 7) for ordering relations
- Connects to goals and causality (Chapters 15, 28) through functionality
- Foundation for spatial and temporal reasoning in later chapters

12.4 Technical Sophistication:

- **Composite Entity Integration**: Scales as sophisticated composite entities with components, orderings, and structural relationships
- **Partial Ordering Theory**: Complete mathematical treatment with antireflexive, antisymmetric, and transitive properties
- **Flexible Scale Relations**: Systematic handling of entities at points vs. points themselves through onScale framework
- **Scale Function Theory**: Monotone-increasing functions preserving ordering relationships across scale mappings

- **Qualitative Structure**: Hi/Md/Lo regional framework enabling approximate and qualitative reasoning

12.5 Complexity Distribution:

- Simple: 14 axioms (basic properties, regional constraints, type definitions)
- Moderate: 17 axioms (multi-condition definitions, scale relations, construction methods)
- Complex: 7 axioms (fundamental scale definition, ordering preservation, composite scale construction)

12.6 Conceptual Importance:

This chapter provides **essential infrastructure** for qualitative and quantitative reasoning in commonsense domains. Scales enable systematic representation of comparative judgments, measurement, and approximate reasoning crucial for practical intelligence. The qualitative Hi/Md/Lo framework captures human-like approximate reasoning, while the mathematical foundation ensures consistent reasoning about ordering relationships. The connection between scalar properties and functionality provides foundation for goal-oriented reasoning.

12.7 Cross-Chapter Connections:

- **Chapter 10 (Composite Entities)**: Scales implemented as composite entities with structural properties
- **Chapter 7 (Substitution)**: Scale ordering relations use substitution framework
- **Chapter 13 (Arithmetic)**: Numeric scales provide quantitative foundation
- **Chapter 14 (Change)**: Vertical scales enable increase/decrease reasoning
- **Chapter 15 (Causality)**: Scale relationships support causal reasoning about degrees
- **Chapter 18 (Space)**: Spatial scales for location and distance reasoning
- **Psychology Chapters**: Psychological scales (intelligence, emotion intensity) throughout theories

12.8 Applications Mentioned:

- **Physical Scales**: height, temperature, spatial location, weight, speed
- **Abstract Scales**: difficulty, importance, prices, probabilities, preferences
- **Qualitative Physics**: positive/zero/negative value reasoning
- **Natural Language**: Comparative (“taller”) and absolute (“tall”) adjective interpretation
- **Goal Reasoning**: Hi region properties as defeasibly goal-relevant
- **Measurement Theory**: Systematic treatment of quantitative relationships

12.9 Notable Design Decisions:

- **Composite Entity Foundation**: Scales as structured entities rather than simple orderings
- **Flexible Point Treatment**: Entities can be at points or be points themselves
- **Regional Structure**: Systematic Hi/Md/Lo decomposition for qualitative reasoning
- **Functionality Connection**: Hi region properties linked to goal achievement
- **Construction Methods**: Multiple approaches for building scales from orderings and structures

12.10 Theoretical Significance:

Chapter 12 establishes the **measurement and comparison infrastructure** essential for sophisticated commonsense reasoning. The scale framework enables systematic treatment of degrees, comparisons, and approximate judgments that pervade human reasoning. The qualitative structure captures essential features of human cognitive processing, while the mathematical foundation ensures consistent reasoning about ordering relationships. This infrastructure is crucial for psychological theories involving degrees of belief, emotion intensity, and comparative judgments.

12.11 Unique Contributions:

12.11.1 Qualitative Scale Structure:

Systematic Hi/Md/Lo regional framework enabling human-like approximate reasoning while maintaining mathematical rigor.

12.11.2 Unified Measurement Theory:

Comprehensive framework integrating quantitative measurement with qualitative reasoning for practical intelligence applications.

12.11.3 Composite Entity Integration:

Novel treatment of scales as structured composite entities enabling rich reasoning about measurement relationships.

12.11.4 Functionality Connection:

Innovative link between scalar properties and goal relevance, providing foundation for preference and value reasoning.

12.11.5 Construction Framework:

Systematic methods for building complex scales from simpler orderings and structural relationships.

This chapter provides the **essential measurement infrastructure** that enables sophisticated reasoning about degrees, comparisons, and approximate relationships throughout commonsense psychology and practical reasoning systems.

Chapter 13

Chapter 13: Arithmetic

- **43 axioms total** covering Peano arithmetic, rational numbers, measures, and half-orders of magnitude
- **4 main sections**: Integers, Rational Numbers, Measures and Proportions, and Half-Orders of Magnitude
- **All background theory** - mathematical foundations

13.1 Key Features Identified:

1. **Peano Arithmetic Foundation**:
   - Axioms 13.1-13.4: Base case (0), successor existence, constraints, and constants 0-10
   - Axioms 13.5-13.7: Positive integers, zero has no predecessor, successor uniqueness
   - Axioms 13.8-13.12: Addition (base case, recursive step, closure, associativity, commutativity)
   - Axioms 13.13-13.19: Multiplication (base case, recursive step, identity, properties, distributive law)
   - Axioms 13.20-13.24: Ordering relations (lt, leq, gt, geq) and number type
2. **Rational Numbers**:
   - Axiom 13.25: Fraction existence for nonnegative numerator and positive denominator
   - Axiom 13.26: Fraction equality using cross-multiplication
   - Axiom 13.27: Integer embedding (n = n/1)
   - Axioms 13.28-13.29: Extension of ordering and multiplication to fractions
   - Axioms 13.30-13.31: Type constraints for fractions
3. **Numeric Scales and Measures**:
   - Axioms 13.32-13.33: Nonnegative numeric scales with lt ordering
   - Axiom 13.34: Measures as monotone-increasing functions mapping bottom to zero
   - Axiom 13.35: Cardinality as measure on set-subset scales
   - Axioms 13.36-13.37: Proportions as ratios f = m(x)/m(y)

- Axioms 13.38-13.40: Identity functions and measure-at relations

4. **Half-Orders of Magnitude**:
   - Axioms 13.41-13.43: Qualitative reasoning framework where entities are "same HOM" if their proportion squared ≤ 10
   - Captures commonsense approximate reasoning ("about", "approximately", "several")
5. **Notable Design Decisions**:
   - **Predicate-based**: Uses sum(n1,n2,n3) instead of n1+n2=n3 for consistency
   - **Constants**: Only constants in the entire book are integers 0-10
   - **Practical note**: Real systems would use procedural attachment, not theorem proving
   - **Extension ready**: Designed to extend to other number types if needed
6. **Complexity Distribution**:
   - Simple: 24 axioms (basic definitions, properties)
   - Moderate: 19 axioms (recursive definitions, multi-condition axioms)
   - Complex: 0 axioms (well-structured mathematical content)

13.2 Conceptual Importance:

This chapter provides essential mathematical infrastructure for: - **Quantitative Reasoning**: Foundation for all numeric comparisons and calculations - **Scale Theory**: Integration with Chapter 12's scale framework - **Measurement**: Formal treatment of measurement functions and proportional relationships - **Qualitative Reasoning**: Half-orders of magnitude for approximate reasoning - **Psychology Applications**: Numeric foundations for psychological scales and measurements

13.3 Cross-Chapter Connections:

- **Chapter 12 (Scales)**: Provides numeric scales and measurement functions
- **Chapter 9 (Functions)**: Uses function theory for measures
- **Psychology Chapters**: Provides quantitative foundations for psychological reasoning
- **Natural Language**: Half-orders of magnitude capture linguistic approximation

13.4 Technical Sophistication:

- **Complete Peano Arithmetic**: Rigorous implementation with base case, successor function, and inductive properties
- **Rational Number System**: Systematic extension from integers to fractions with cross-multiplication equality
- **Measure Theory**: Formal treatment of measurement functions as monotone-increasing mappings
- **Proportional Reasoning**: Systematic framework for ratios and proportional relationships

- **Half-Orders of Magnitude**: Novel qualitative reasoning framework for approximate numeric judgments

13.5 Complexity Distribution:

- Simple: 24 axioms (basic definitions, properties, type constraints)
- Moderate: 19 axioms (recursive definitions, multi-condition axioms, complex relationships)
- Complex: 0 axioms (well-structured mathematical content with clear logical organization)

13.6 Conceptual Importance:

This chapter provides the **fundamental quantitative foundation** for all numeric reasoning in commonsense domains. The arithmetic framework enables systematic reasoning about quantities, measurements, and proportions essential for practical intelligence. The half-orders of magnitude framework captures human-like approximate reasoning, while the measure theory provides rigorous foundation for quantitative relationships. This infrastructure supports both precise calculation and qualitative approximation crucial for psychological and practical reasoning.

13.7 Cross-Chapter Connections:

- **Chapter 12 (Scales)**: Provides numeric scales and measurement functions for qualitative reasoning
- **Chapter 9 (Functions)**: Uses function theory for systematic treatment of measures and mappings
- **Chapter 6 (Sets)**: Cardinality as canonical measure connecting set theory to arithmetic
- **Chapter 14 (Change)**: Quantitative changes and increases/decreases on numeric scales
- **Psychology Chapters**: Quantitative foundations for psychological scales, probabilities, and measurements
- **Planning Chapters**: Resource counting, time estimation, and quantitative constraint reasoning

13.8 Applications Mentioned:

- **Physical Measurements**: size, time, money, weight with half-order approximations
- **Cognitive Judgments**: linguistic approximations (“about,” “approximately,” “several,” “nearly”)
- **Practical Reasoning**: scale-dependent interaction modes (carrying capacity, meeting duration)
- **Set Theory**: cardinality as fundamental measure connecting discrete and continuous reasoning
- **Economic Reasoning**: monetary values, costs, and proportional relationships

- **Time Management**: duration estimation and temporal quantification

13.9 Notable Design Decisions:

- **Predicate-Based Arithmetic**: sum(n1,n2,n3) instead of n1+n2=n3 for logical consistency
- **Limited Constants**: Only constants in entire book are integers 0-10 for practical reasoning
- **Procedural Attachment**: Recognition that real systems need computational arithmetic, not theorem proving
- **Extension Ready**: Framework designed to accommodate additional number types as needed
- **Qualitative Integration**: Half-orders bridge precise arithmetic and approximate human reasoning

13.10 Theoretical Significance:

Chapter 13 establishes the **quantitative reasoning foundation** essential for sophisticated commonsense intelligence. The arithmetic framework enables precise calculation while the half-orders of magnitude capture human-like approximate reasoning crucial for practical decision-making. The measure theory provides rigorous foundation for quantitative relationships in physical and abstract domains, while the proportional reasoning framework supports comparative and relative judgments essential for psychological reasoning.

13.11 Unique Contributions:

13.11.1 Half-Orders of Magnitude Framework:

Innovative qualitative reasoning system where entities are "same HOM" if proportion squared ≤ 10, capturing commonsense approximation.

13.11.2 Integrated Quantitative System:

Comprehensive framework spanning precise arithmetic to approximate reasoning, supporting both computational and cognitive applications.

13.11.3 Measure Theory Integration:

Systematic treatment of measurement functions connecting abstract mathematical concepts to practical measurement scenarios.

13.11.4 Proportional Reasoning Framework:

Rigorous foundation for ratio-based reasoning essential for comparative judgments and relative assessments.

13.11.5 Cognitive Approximation Bridge:

Novel connection between precise mathematical reasoning and human-like approximate quantitative judgments.

This chapter provides the **essential quantitative infrastructure** enabling sophisticated numeric reasoning across precise calculation and approximate judgment, supporting both the mathematical foundations and psychological theories throughout the comprehensive commonsense reasoning system.

Chapter 14

Chapter 14: Change of state

- **13 axioms total** covering change of state and derived concepts
- **2 main sections**: The change Predicate and Predicates Derived from change
- **All background theory** - foundational concepts for temporal and causal reasoning

14.1 Key Features Identified:

1. **Basic Change Framework**:
   - Axiom 14.1: Change arguments must be eventualities
   - Axiom 14.2: Changes must involve a common entity (prevents unrelated state changes)
   - Axiom 14.3: Change is defeasibly transitive (with etc conditions)
2. **Inconsistency and Cyclical Change**:
   - Axiom 14.4: Complex axiom handling inconsistency - if states aren't inconsistent, change must go through an inconsistent intermediate state
   - Axiom 14.5: Defeasible inference that start and end states are inconsistent (since change isn't normally cyclic)
3. **Derived Change Predicates**:
   - Axiom 14.6: changeIn - change in properties of an entity
   - Axiom 14.7: changeFrom - change out of a state (ensures no same-type state exists after)
   - Axiom 14.8: changeTo - change into a state (ensures no same-type state existed before)
4. **Movement and Vertical Scales**:
   - Axiom 14.9: move - change from being at one location to another
   - Axioms 14.10-14.11: Vertical scales (numeric scales are vertical; vertical arguments must be scales)
   - Axioms 14.12-14.13: increase/decrease as movement up/down vertical scales
5. **Technical Sophistication**:
   - **Reified Predicates**: Heavy use of primed predicates (change', at', etc.)
   - **Substitution**: Uses substitution framework from Chapter 7 for type/token distinctions

- **Generation**: Uses gen relation for eventuality relationships
- **Inconsistency**: Sophisticated handling of when states conflict

6. **Complexity Distribution**:
   - Simple: 6 axioms (basic constraints, type requirements)
   - Moderate: 6 axioms (derived predicate definitions)
   - Complex: 1 axiom (14.4 - change inconsistency requirement)

14.2 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Temporal Reasoning**: Foundation for understanding processes and events over time - **Causal Reasoning**: Change is prerequisite for causality - **Natural Language**: Verbs of motion and change ("move," "increase," "become") - **Psychological Processes**: Learning, growth, adaptation all involve change - **Physical Processes**: Motion, transformation, development

14.3 Cross-Chapter Connections:

- **Chapter 7 (Substitution)**: Uses subst for type/token relationships
- **Chapter 8 (Logic Reified)**: Uses inconsistent and gen predicates
- **Chapter 10 (Composite Entities)**: Uses at relation for location
- **Chapter 12 (Scales)**: Uses lts and vertical scale concepts
- **Chapter 5 (Eventualities)**: Fundamental eventuality framework

14.4 Applications Mentioned:

- **Physical**: Moving, growing, opening/closing doors
- **Cognitive**: Learning (not knowing $\rightarrow$ knowing)
- **Social**: Changing relationships, roles
- **Measurement**: Changes in quantities on scales
- **Spatial**: Movement from place to place

14.5 Technical Sophistication:

- **Reified Change Framework**: Change as fundamental predicate relating eventualities with sophisticated argument constraints
- **Inconsistency Management**: Complex logic for handling when states conflict and intermediate inconsistent states
- **Derived Predicate System**: Systematic family of change predicates (changeIn, changeFrom, changeTo) with precise semantics

- **Vertical Scale Integration**: Movement up/down scales for increase/decrease with mathematical precision
- **Defeasible Transitivity**: Change propagation with etc conditions for realistic reasoning

14.6 Complexity Distribution:

- Simple: 6 axioms (basic constraints, type requirements, simple definitions)
- Moderate: 6 axioms (derived predicate definitions with multiple conditions)
- Complex: 1 axiom (inconsistency requirement 14.4 with sophisticated logical structure)

14.7 Conceptual Importance:

This chapter provides the **fundamental temporal infrastructure** for reasoning about processes, events, and transformations in commonsense domains. Change is the foundation for causality, learning, development, and all dynamic processes essential to psychological and physical reasoning. The framework enables systematic reasoning about state transitions while handling the complexities of inconsistency and temporal relationships. This infrastructure is crucial for representing the dynamic nature of mental states and processes.

14.8 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Fundamental eventuality framework underlying all change relationships
- **Chapter 7 (Substitution)**: Type/token distinctions crucial for change reasoning
- **Chapter 8 (Logic Reified)**: Inconsistency predicates and generation relationships
- **Chapter 10 (Composite Entities)**: At relation for location-based change
- **Chapter 12 (Scales)**: Vertical scales enabling increase/decrease reasoning
- **Chapter 15 (Causality)**: Change as prerequisite for causal relationships
- **Chapter 16 (Time)**: Temporal framework for sequencing changes
- **Psychology Chapters**: Learning, emotion, and cognitive changes throughout theories

14.9 Applications Mentioned:

- **Physical Processes**: movement, growth, opening/closing, transformation
- **Cognitive Development**: learning transitions (not knowing $\rightarrow$ knowing)
- **Social Dynamics**: relationship changes, role transitions
- **Measurement Changes**: quantitative increases/decreases on scales
- **Spatial Movement**: location changes and navigation
- **Psychological Processes**: belief revision, emotion changes, goal shifts

14.10 Notable Design Decisions:

- **Fundamental Treatment**: Change as undefined predicate too basic for formal definition
- **Eventuality Foundation**: Changes relate eventualities rather than simple entities
- **Defeasible Framework**: Exception handling through etc predicates for realistic reasoning
- **Vertical Metaphor**: Systematic "up/down" treatment for numeric and abstract scales
- **Inconsistency Requirement**: Sophisticated handling of when states conflict during transitions

14.11 Theoretical Significance:

Chapter 14 establishes the **essential dynamic foundation** for commonsense reasoning about processes and transformations. Change is fundamental to causality, learning, and all temporal reasoning essential for psychological intelligence. The framework provides systematic method for reasoning about state transitions while handling the logical complexities of inconsistency and temporal relationships. This infrastructure enables the sophisticated reasoning about mental processes, learning, and development central to commonsense psychology.

14.12 Unique Contributions:

14.12.1 Reified Change Framework:

Systematic treatment of change as fundamental predicate relating eventualities, enabling meta-reasoning about transformations.

14.12.2 Inconsistency Management System:

Sophisticated logic for handling conflicting states and intermediate inconsistent conditions during transitions.

14.12.3 Derived Change Taxonomy:

Comprehensive family of change predicates (changeIn, changeFrom, changeTo) with precise logical semantics.

14.12.4 Vertical Scale Integration:

Innovative connection between abstract change concepts and concrete scale-based increase/decrease reasoning.

14.12.5 Defeasible Transition Logic:

Realistic framework for change reasoning that accommodates exceptions and context-dependent transitions.

This chapter provides the **fundamental dynamic infrastructure** essential for reasoning about the temporal processes, transformations, and developments that characterize both physical systems and psychological states throughout sophisticated commonsense reasoning.

Chapter 15

Chapter 15: Causality

- **36 axioms total** covering causality, agency, and related causal concepts
- **6 main sections**: Causal Complexes, Agents and Agenthood, Other Causal Predicates, Ability, Executability, and Difficulty
- **All background theory** - foundational concepts for understanding causation and agency in commonsense reasoning

15.1 Key Features Identified:

1. **Causal Complex Framework**:
   - Axiom 15.1: Causal complex arguments must be eventualities or sets of eventualities
   - Axiom 15.2: Member relevance - removing any element breaks the causal complex
   - Axiom 15.3: Causes must be members of causal complexes
   - Axiom 15.4: Defeasible transitivity of causation (with etc conditions)
2. **Agent Theory and Willing**:
   - Axiom 15.5: Complex distinction between eventuality causation (cause0) and agent causation (via willing)
   - Axiom 15.6: Agents defeasibly have causal capabilities
   - Axiom 15.7-15.8: Agent case roles and action definitions
   - Axioms 15.9-15.12: Linguistic case roles (object, instrument, source, terminus) defined causally
3. **Extended Causal Vocabulary**:
   - Axiom 15.14: causallyInvolved - broader than direct causation
   - Axioms 15.15-15.17: enable/enable0 - preconditions vs. causes
   - Axioms 15.18-15.19: allow/prevent - permissive vs. blocking relations
   - Axioms 15.20-15.23: partiallyCause, tcause, tcauseq - weaker causal notions
4. **Ability and Control**:
   - Axiom 15.24: Complex definition of events beyond an agent's control

- Axiom 15.25: able - possibility when uncontrollable factors cooperate
- Axiom 15.26: ability as reified state of being able

5. **Executability Theory**:
   - Axiom 15.27: dcause - direct causation without intermediates
   - Axioms 15.28-15.29: Willing as source of direct causation for agents
   - Axiom 15.30: enabled - when causal complex preconditions hold at time t
   - Axiom 15.31: Complex recursive definition of executable actions
6. **Difficulty Framework**:
   - Axiom 15.32: Type constraints for difficult actions
   - Axiom 15.33: difficultiesWith - obstructions that prevent action success
   - Axioms 15.34-15.35: Difficulty as high position on obstruction-consistent scales

15.2 Technical Sophistication:

- **Reified Predicates**: Extensive use of primed predicates (will', cause', changeIn', not', and', at', change', difficult', able', dcause')
- **Defeasible Reasoning**: Two axioms use (etc) for non-monotonic inference
- **Recursive Definitions**: objectOf and executable defined recursively
- **Modal Integration**: References possibility, likelihood from Chapter 20
- **Temporal Integration**: Uses atTime from Chapter 16

15.3 Complexity Distribution:

- Simple: 11 axioms (basic constraints, type requirements, simple implications)
- Moderate: 12 axioms (standard definitional equivalences)
- Complex: 13 axioms (deep nesting, multiple quantifiers, recursive structure)

15.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Planning and Reasoning**: Understanding what agents can cause and control - **Explanation**: Distinguishing causes from mere conditions or correlations - **Natural Language**: Verbs of causation, agency, and capability - **Psychological Processes**: Intention, action, skill, effort, success/failure - **Practical Reasoning**: Ability assessment, planning, execution

15.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Fundamental eventuality and real existence framework
- **Chapter 6 (Sets)**: Set operations for causal complexes
- **Chapter 8 (Logic Reified)**: Uses and', not', imply predicates

- **Chapter 12 (Scales)**: Uses subsetConsistent for difficulty ordering
- **Chapter 16 (Time)**: Uses atTime for temporal constraints
- **Chapter 20 (Modality)**: Uses possible, likely, likelihood predicates
- **Chapters 19/31**: References will predicate for agent volition

15.6 Applications Mentioned:

- **Physical Causation**: Light switches, car engines, falling objects
- **Agency**: Human and robotic action, planning, skill execution
- **Explanation**: Distinguishing causes from background conditions
- **Planning**: Executability assessment, obstacle identification
- **Capability**: Ability vs. opportunity, skill vs. circumstance

15.7 Notable Design Decisions:

- **Two-Level Causation**: Distinguishes precise causal complexes from practical defeasible causes
- **Agent Integration**: Seamlessly integrates physical and agential causation through willing
- **Linguistic Grounding**: Defines case roles in terms of causal structure
- **Probabilistic Extension**: Includes tcause for uncertain/partial causation
- **Practical Focus**: Emphasizes executability and difficulty for real-world reasoning
- **Modal Sophistication**: Integrates with possibility and likelihood theories

15.8 Theoretical Significance:

Chapter 15 represents one of the most sophisticated treatments of causality in the book, providing both the mathematical precision needed for formal reasoning and the practical concepts required for commonsense psychology. The distinction between causal complexes and causes mirrors the difference between scientific and folk psychological explanations, while the integration of agency through willing provides a bridge between physical and intentional causation.

Chapter 16

Chapter 16: Time

- **80 axioms total** covering temporal ontology, eventuality-time relations, temporal ordering, durations, periodicity, and rates
- **6 main sections**: Temporal Topology, Eventuality-Time Relations, Temporal Ordering, Durations, Periodicity, and Rates/Frequency
- **All background theory** - foundational temporal infrastructure for commonsense reasoning

16.1 Key Features Identified:

1. **Temporal Ontology Foundation**:
   - Axioms 16.1-16.2: Basic types - instants and intervals as temporal entities
   - Axioms 16.3-16.5: Argument constraints for begins, ends, insideTime relations
   - Axioms 16.6-16.7: Instant beginning/end identity (instant is its own boundary)
   - Axioms 16.8-16.10: Interval definitions (intervalBetween, posInfInterval, properInterval)
2. **Temporal Sequences**:
   - Axiom 16.11: Complex definition of temporal sequences as nonoverlapping temporal entities with gaps
   - Axioms 16.12-16.14: first, last, and successiveElts for temporal sequences
   - Axioms 16.15-16.17: Temporal sequences as scales with before ordering and convex hulls
   - Axioms 16.18-16.21: Extension of begins/ends/insideTime to temporal sequences
3. **Eventuality-Time Integration**:
   - Axiom 16.22: atTime argument constraints (eventuality at instant)
   - Axiom 16.23: during definition (eventuality throughout proper interval)
   - Axiom 16.24: Complex timeSpanOf definition covering instants, intervals, and temporal sequences
   - Axioms 16.25-16.30: Extensions (happensIn, temporal boundaries, infinite intervals)
4. **Temporal Ordering and Causality**:
   - Axioms 16.36-16.41: before relation properties (antireflexive, antisymmetric, transitive)

- Axioms 16.42-16.45: Allen's interval algebra (intMeets, intOverlap, intFinishes, intDuring)
- Axioms 16.51-16.53: Temporal constraints on change, causation, and enablement
- Axioms 16.54-16.55: Causal complex effects and defeasible causal modus ponens

5. **Duration and Measurement**:
   - Axioms 16.57-16.62: sameDuration relation and temporal units
   - Axiom 16.63: Complex concatenation definition for intervals
   - Axioms 16.64-16.69: durationOf predicate for all temporal entity types
   - Axioms 16.70-16.72: Duration-based scales and ordering
6. **Periodicity and Rates**:
   - Axioms 16.73-16.75: Periodic and roughly periodic temporal sequences
   - Axioms 16.76-16.78: Complex rate definitions for events per time unit
   - Axioms 16.79-16.80: Rate scales and frequency as high-rate positioning

16.2 Technical Sophistication:

- **OWL-Time Based**: Condensed from OWL-Time ontology with commonsense modifications
- **Allen's Interval Algebra**: Implements subset of Allen's 13 interval relations
- **Multi-Level Coercion**: Automatic coercion from eventualities to temporal entities
- **Duration Without Numbers**: Duration based on temporal units and concatenation, not real numbers
- **Scale Integration**: Temporal entities as scales with before ordering

16.3 Complexity Distribution:

- Simple: ~25 axioms (basic constraints, type requirements, simple relations)
- Moderate: ~35 axioms (standard definitions, temporal extensions)
- Complex: ~20 axioms (multi-case definitions, nested quantification, interval algebra)

16.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Temporal Reasoning**: Foundation for understanding when events occur and their temporal relationships - **Causal Reasoning**: Temporal constraints preventing effects before causes - **Natural Language**: Temporal expressions, tense, aspect, temporal adverbials - **Planning and Scheduling**: Duration estimation, temporal coordination, deadlines - **Narrative Understanding**: Sequence, simultaneity, temporal progression

16.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Fundamental eventuality framework for temporal attribution
- **Chapter 9 (Sequences)**: Sequence operations extended to temporal sequences

- **Chapter 12 (Scales)**: Time as scale, duration scales, rate scales, half-orders of magnitude
- **Chapter 14 (Change)**: Temporal constraints on state changes
- **Chapter 15 (Causality)**: Temporal ordering constraints on causal relations
- **Chapter 18 (Space)**: Parallel at relation structure (time as metaphorical location)

16.6 Applications Mentioned:

- **Physical Events**: Motion, collision, process duration
- **Scheduling**: Meeting times, deadlines, coordination
- **Measurement**: Driving speed (60 mph), meeting frequency (3 per month)
- **Periodicity**: Regular events, roughly periodic patterns
- **Causal Reasoning**: Cause-effect temporal sequences, enabling conditions

16.7 Notable Design Decisions:

- **No Linear Time Assumption**: Allows for branching or partial temporal orderings
- **Instant vs. Interval Neutrality**: Doesn't assume intervals are composed of instants
- **Unit-Based Duration**: Avoids real number mappings in favor of temporal unit comparisons
- **Defeasible Causation**: Causal modus ponens includes (etc) condition for exceptions
- **Eventuality Coercion**: Automatic handling of eventualities in temporal contexts
- **Rate Generalization**: Sophisticated rate conversion between different temporal units

16.8 Theoretical Significance:

Chapter 16 represents the most comprehensive temporal theory in the book, providing both the mathematical precision needed for formal temporal reasoning and the flexibility required for commonsense temporal expressions. The integration with causality and change creates a unified framework where temporal, causal, and change relationships are mutually constraining. The duration theory based on temporal units rather than real numbers reflects how humans actually think about time measurement, while the rate and periodicity frameworks enable reasoning about temporal patterns in everyday life.

The chapter's treatment of temporal sequences as scales enables sophisticated qualitative temporal reasoning, while the Allen interval algebra provides precise tools for temporal interval relationships. The careful separation of instants, intervals, and temporal sequences with systematic coercion rules creates a robust foundation for representing the full complexity of temporal phenomena in commonsense psychology.

Chapter 17

Chapter 17: Event structure

- **16 axioms total** covering event structure, subevents, sequences, conditionals, and iterations
- **3 main sections**: Events and Subevents, Event Sequences and Conditionals, Iterations
- **All background theory** - foundational concepts for structured event representation and control flow

17.1 Key Features Identified:

1. **Basic Event and Subevent Framework**:
   - Axioms 17.1-17.2: Subevent relation properties (antisymmetric, transitive)
   - Axiom 17.3: Complex event definition - events involve change directly, through generation, or via subevents
   - Axioms 17.4-17.5: Type constraints (events are eventualities, subevents relate events)
2. **Event Aggregation and Sequences**:
   - Axioms 17.6-17.7: Conjunction creates events - and' of eventualities with at least one event yields an event
   - Axiom 17.8: Event sequence definition - temporally ordered events with reified conjunction
   - Axiom 17.9: Sequence subevent properties - components are subevents of the sequence
3. **Conditional Events**:
   - Axiom 17.10: Complex conditional definition - implication where condition holds at event beginning
   - Axiom 17.11: Conditional subevent properties - consequent is subevent of conditional
4. **Iterative Control Structures**:
   - Axiom 17.12: Recursive iteration definition - pure iteration with no termination condition
   - Axiom 17.13: WhileDo recursive definition - iteration with continuing condition check
   - Axiom 17.14: RepeatUntil recursive definition - iteration with termination condition check
   - Axiom 17.15: ForAllOfSeq recursive definition - iteration over sequence elements

5. **Composite Entity Integration**:
    - Axiom 17.16: Complex characterization of events as composite entities with subevents as components

17.2 Technical Sophistication:

- **Programming Language Metaphor**: World viewed as computer executing its own history
- **Recursive Definitions**: Four axioms use recursive structure for iteration constructs
- **Reified Predicates**: Extensive use of change', and', imply', eventSequence', event', subevent'
- **Control Flow Structures**: Systematic treatment of sequence, conditional, iteration paralleling programming languages
- **Composite Entity Framework**: Integration with Chapter 10's composite entity theory

17.3 Complexity Distribution:

- Simple: 5 axioms (basic constraints, type requirements, simple implications)
- Moderate: 3 axioms (standard definitional equivalences)
- Complex: 8 axioms (recursive definitions, multiple quantifiers, sophisticated logical structure)

17.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Narrative Understanding**: Complex events with internal structure and temporal flow - **Process Modeling**: Iterative and conditional processes in natural and social domains - **Action Planning**: Structured decomposition of complex goals into subevent sequences - **Natural Language**: Event descriptions with embedded control structures - **Cognitive Modeling**: Mental representation of complex, structured activities

17.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Fundamental eventuality framework and generation relation
- **Chapter 7 (Substitution)**: Uses instance and subst predicates for type instantiation
- **Chapter 8 (Logic Reified)**: Uses and', imply' for reified logical operations
- **Chapter 10 (Composite Entities)**: Events as composite entities with subevent components
- **Chapter 14 (Change)**: Events fundamentally involve change' relations
- **Chapter 16 (Time)**: Uses atTime, begins, beforeOrMeets for temporal constraints

17.6 Applications Mentioned:

- **Natural Processes**: Sun rising/setting, hourglass sand falling, iterative natural phenomena
- **Programming**: Control structures (sequence, conditional, while/repeat loops, for-each)

- **Narrative**: "Day was warm and Pat jogged" - conjunctive events
- **Physical Processes**: "Object rolls into water; if dense, it sinks" - conditional events
- **Computational Metaphor**: World as computer executing its own temporal history

17.7 Notable Design Decisions:

- **Events Require Change**: Fundamental commitment that events involve state changes
- **Programming Language Structure**: Deliberate parallel with control flow constructs
- **Recursive Iteration**: Iterations defined recursively rather than as primitive loops
- **Conditional Events Allowed**: Rejecting factoring out implications for ontological richness
- **Composite Entity Integration**: Events as structured wholes with subevent parts
- **Minimal Structure**: At least two elements required for iterations (no trivial cases)

17.8 Theoretical Significance:

Chapter 17 represents a sophisticated fusion of temporal logic, programming language theory, and mereological analysis applied to events. The decision to allow conditional events as genuine events (rather than factoring out the conditionals) reflects a commitment to ontological richness that supports natural language understanding and narrative comprehension.

The recursive definitions of iteration constructs provide a mathematically precise foundation for representing repetitive processes, while the integration with composite entity theory enables reasoning about the internal structure of complex events. The programming language metaphor is particularly powerful, suggesting that natural processes and human activities can be understood using the same structural principles that govern computational control flow.

This framework enables sophisticated reasoning about complex, structured activities that unfold over time with internal dependencies, conditions, and repetitive patterns - essential for understanding everything from natural processes to human goal-directed behavior.

Chapter 18

Chapter 18: Space

- **11 axioms total** covering spatial systems, distance relations, location, and spatial analogies
- **2 main sections**: Spatial Systems and Distance, Location
- **All background theory** - minimal spatial foundation focused on properties used in spatial analogies

18.1 Key Features Identified:

1. **Spatial Metaphor Philosophy**:
   - **Core Insight**: Abstract domains understood through spatial analogies, but only topological properties transfer
   - **Design Decision**: Develop theories of abstract properties (composite entities, scales, change) rather than rich spatial theory
   - **Minimal Spatial Vocabulary**: Only distance, location, and nearness - no color, hardness, precise measurement
2. **Spatial System Framework**:
   - Axiom 18.1: Complex definition of spatial systems as composite entities with physical object components and distance relations
   - Axiom 18.2: Distance argument constraints (non-negative integers, spatial units, system components)
3. **Mathematical Distance Properties**:
   - Axiom 18.3: Distance reflexivity (entity to itself is zero)
   - Axiom 18.4: Distance symmetry (d(x,y) = d(y,x))
   - Axiom 18.5: Triangle inequality (direct path $\leq$ sum of indirect paths)
4. **Nearness as Scale-Based Concept**:
   - Axiom 18.6: Shorter distance definition (comparative distance relation)
   - Axiom 18.7: Complex nearness scale definition (reverse of distance scale so shorter = Hi region)

- Axiom 18.8: Near definition (distance in Hi region of nearness scale)

5. **Location Integration**:
   - Axiom 18.9: Spatial systems as grounds for the at relation
   - Axiom 18.10: atLoc as specialized at relation for spatial systems
   - Axiom 18.11: Proximity inheritance (entities near iff their locations near)

18.2 Technical Sophistication:

- **Composite Entity Foundation**: Built on Chapter 10's composite entity framework
- **Scale Theory Integration**: Nearness uses reversed distance scales from Chapter 12
- **Mathematical Rigor**: Standard metric space properties for distance relations
- **Figure-Ground Relations**: Integration with general at relation from Chapter 10
- **Minimal Commitment**: Deliberately sparse spatial vocabulary focusing on analogy-relevant properties

18.3 Complexity Distribution:

- Simple: 5 axioms (basic constraints, type requirements, simple definitions)
- Moderate: 3 axioms (mathematical properties, location relations)
- Complex: 3 axioms (spatial system definition, nearness scale, near definition)

18.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Spatial Analogies**: Foundation for understanding abstract domains through spatial metaphors - **Physical Object Reasoning**: Basic spatial relationships for commonsense physics - **Location-Based Reasoning**: Where events happen, where agents are located - **Natural Language**: Spatial prepositions, locative expressions, proximity judgments - **Cognitive Modeling**: Spatial cognition as basis for abstract thought

18.5 Cross-Chapter Connections:

- **Chapter 10 (Composite Entities)**: Spatial systems as composite entities, at relation, ground concept
- **Chapter 12 (Scales)**: Nearness scales, Hi regions, scale reversal operations
- **Chapter 13 (Arithmetic)**: Distance measurements, triangle inequality, sum operations
- **Chapter 16 (Time)**: Parallel between atTime and atLoc relations
- **Future Chapters**: Spatial location for agents, events, and psychological states

18.6 Applications Mentioned:

- **Physical Location**: "Jill is at her desk", "typing at her desk"
- **Event Location**: Teleconference occurring at specific location
- **Spatial Analogies**: "in politics", "in trouble", "move debate from politics to religion"
- **Proximity Judgments**: Near vs. far relationships in spatial systems
- **Abstract Domains**: Money moving between bank accounts (spatial metaphor)

18.7 Notable Design Decisions:

- **Topological Focus**: Emphasizes properties commonly transferred in analogies (not color, hardness, exact distance)
- **Minimal Spatial Theory**: Only what's needed for psychology rather than full commonsense physics
- **Scale-Based Nearness**: Near as Hi region of reversed distance scale rather than fixed threshold
- **Integer Distance**: Discrete rather than continuous distance measurements
- **No Geometric Structure**: No angles, shapes, orientations - purely topological
- **Physical Object Primitive**: physobj left undefined to avoid commonsense physics

18.8 Theoretical Significance:

Chapter 18 represents a sophisticated philosophical position about the relationship between spatial and abstract cognition. Rather than building a rich spatial theory and then creating analogical mapping mechanisms, the authors identify the abstract properties most commonly transferred from spatial domains and develop theories of those properties directly.

This "spatial metaphor for understanding space" approach suggests that our spatial cognition itself may be built on more fundamental abstract patterns like composite structure, scalar relationships, and change processes. The minimal spatial vocabulary (distance, location, nearness) focuses on exactly those spatial properties that appear repeatedly in cross-domain analogies.

The integration with scale theory for defining nearness creates a flexible framework where proximity judgments can vary contextually based on the relevant spatial system and scale considerations. This supports both literal spatial reasoning and the metaphorical extensions that pervade natural language and abstract thought.

The chapter's restraint in avoiding geometric detail reflects its focus on commonsense psychology rather than physics or robotics, providing just enough spatial infrastructure to support the cognitive theories that follow while maintaining the book's emphasis on abstract structural relationships over domain-specific details.

Chapter 19

Chapter 19: Persons

- **14 axioms total** covering the basic structure of persons as agents with bodies and minds, perception, and voluntary control
- **3 main sections**: Basic Person Structure, Perception, Voluntary Control
- **Mostly background theory** with one example - foundational concepts for human-specific aspects of cognition

19.1 Key Features Identified:

1. **Basic Person Ontology**:
   - Axiom 19.1: Person is a kind of agent (inheriting causal capabilities)
   - Axiom 19.2: Person is a kind of physical object (spatial existence)
   - Axiom 19.3: Persons have both body and mind (mind-body structure)
   - Axioms 19.4-19.5: Mind as composite entity, body as physical object
2. **Body as Spatial System**:
   - Axiom 19.6: Body is a spatial system with body parts as components
   - Axiom 19.7: Body parts set definition in terms of body components
   - Integration with spatial framework from Chapter 18
3. **Perception Framework**:
   - Axiom 19.8: Perception relates agents to entities external to mind
   - Axiom 19.9: Nearness as enabling condition for perception (spatial constraint)
   - Axioms 19.10-19.11: Sense organs as distinguished subset of body parts
   - Axiom 19.12: Intact sense organs as enabling conditions for perception
4. **Voluntary Control System**:
   - Axiom 19.13: Complex definition of directly controllable body parts via willing and direct causation
   - Axiom 19.14: Example of voluntary arm lifting as illustration

19.2 Technical Sophistication:

- **Mind-Body Integration**: Systematic treatment without committing to specific philosophical positions
- **Composite Entity Framework**: Both minds and bodies as composite entities with internal structure
- **Spatial System Integration**: Bodies as spatial systems connecting to Chapter 18's framework
- **Enabling Conditions**: Multiple enabling conditions (nearness, intact sense organs) for perception
- **Direct Causation**: Voluntary control through will causing direct body part movement

19.3 Complexity Distribution:

- Simple: 7 axioms (basic type constraints, simple implications)
- Moderate: 6 axioms (definitional structures, enabling conditions)
- Complex: 1 axiom (voluntary control with multiple nested quantifiers)

19.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Embodied Cognition**: Foundation for understanding mind-body interaction - **Perception Theory**: Basic framework for sensory input and spatial constraints - **Action Theory**: Voluntary control and direct causation of bodily movement - **Agent Capabilities**: Distinguishing person-specific from general agent properties - **Commonsense Psychology**: Physical basis for mental phenomena

19.5 Cross-Chapter Connections:

- **Chapter 15 (Causality)**: Uses agent, enable, dcause predicates for causal relationships
- **Chapter 18 (Space)**: Bodies as spatial systems with physobj components and nearness relations
- **Chapter 10 (Composite Entities)**: Minds and bodies as composite entities with components
- **Chapter 28 (Goals)**: References intact predicate for functionality (future)
- **Chapter 31 (Plans)**: References will predicate for intention-action conversion (future)

19.6 Applications Mentioned:

- **Perception**: Spatial constraints on what can be perceived, sense organ requirements
- **Voluntary Action**: Direct control of body parts (arm lifting) vs. indirect control (hair movement)
- **Mind-Body Interaction**: Perception and action as interfaces between mental and physical

- **Physical Embodiment**: Persons as both cognitive agents and spatial physical objects

19.7 Notable Design Decisions:

- **Dual Nature**: Persons as both agents and physical objects (avoiding category errors)
- **Mind-Body Neutrality**: Composite entity treatment without philosophical commitment
- **Spatial Embodiment**: Bodies as spatial systems integrating cognitive and physical theories
- **Enabling Conditions**: Multiple necessary conditions for perception (proximity, intact organs)
- **Direct vs. Indirect Control**: Voluntary control limited to subset of body parts
- **Minimal Anatomy**: Avoids detailed anatomical commitments while establishing functional categories

19.8 Theoretical Significance:

Chapter 19 represents a crucial transition point in the book, introducing the first specifically human elements while maintaining the abstract framework developed in earlier chapters. The treatment of persons as both agents and physical objects provides a foundation for embodied cognition without committing to specific philosophical positions about mind-body relationships.

The spatial embedding of bodies and the enabling conditions for perception create a naturalistic framework where cognitive phenomena are constrained by physical realities. The distinction between directly and indirectly controllable body parts provides a foundation for understanding the limits and mechanisms of voluntary action.

The composite entity treatment of both minds and bodies enables systematic analysis of internal structure while maintaining the unified theoretical framework. This supports later developments in cognitive theory by providing the physical and perceptual infrastructure necessary for beliefs, goals, plans, and other psychological phenomena.

The chapter's restraint in avoiding detailed anatomical or neurological commitments reflects its focus on commonsense psychology rather than scientific psychology, providing just enough physical detail to support the cognitive theories that follow while maintaining the book's emphasis on folk psychological understanding.

Chapter 20

Chapter 20: Modality

- **33 axioms total** covering modes of existence including real existence, time, possibility, necessity, and qualitative probability
- **4 main sections**: Rexist and atTime, Positive Modalities, Possibility and Necessity, Likelihood (Qualitative Probability)
- **All background theory** - foundational modal framework for reasoning about different modes of existence

20.1 Key Features Identified:

1. **Real Existence and Time Framework**:
   - Axiom Schema 20.1: Unprimed predicates equivalent to reified predicates with Rexist
   - Axiom Schema 20.2: Temporal arguments equivalent to atTime relations
   - Axioms 20.4-20.6: Now as unique temporal anchor connecting Rexist to atTime
   - Foundation for relating atemporal logic to temporal reality
2. **Positive Modalities**:
   - Axiom 20.7: PosMod as cover term for modalities supporting inference
   - Axiom 20.8: Rexist as exemplar positive modality
   - Axiom 20.9: Modus ponens preservation within positive modalities
   - Framework enabling systematic treatment of inference-supporting modes
3. **Constraint-Relative Possibility and Necessity**:
   - Axioms 20.10-20.11: Possibility as absence of constraint-implied negation
   - Axioms 20.13-20.14: Necessity as constraint-implied conclusion
   - Axioms 20.12,20.15: Both as positive modalities
   - Axioms 20.16-20.17: Duality relationships and impossibility definition
4. **Qualitative Likelihood Theory**:
   - Axiom 20.18-20.19: Likelihood scales with partial ordering (not total numerical)
   - Axioms 20.20-20.21: AlsoRequired framework linking likelihood to minimal assumptions

- Axioms 20.22-20.26: Conjunction/disjunction operations and scale extremes
- Axioms 20.27-20.33: Likelihood as scale-based concept with Hi region definition

5. **Scale Integration**:
   - Likelihood scales as partially ordered (not total like probability)
   - Integration with Chapter 12's scale framework (Hi regions, vertical scales)
   - Grounds for likelihood scales enabling "at" relations
   - Comparative likelihood through scale ordering

20.2 Technical Sophistication:

- **Axiom Schemas**: Two schemas relating different predicate styles and temporal representations
- **Modal Logic Foundation**: Systematic treatment of alethic modalities with constraint relativity
- **Qualitative Probability**: Non-numerical alternative to mathematical probability theory
- **Scale-Based Modality**: Likelihood defined through scale positions rather than numbers
- **Positive Modality Theory**: Unified framework for inference-preserving modalities

20.3 Complexity Distribution:

- Simple: 20 axioms (basic constraints, type requirements, simple definitions)
- Moderate: 9 axioms (modal definitions, scale operations)
- Complex: 4 axioms (axiom schemas, likelihood scale definition, constraint-based definitions)

20.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Belief and Knowledge**: Modal framework for epistemic states and uncertainty - **Planning and Decision**: Possibility and likelihood for action selection - **Natural Language**: Modal expressions ("possible", "likely", "necessary") - **Temporal Reasoning**: Connecting atemporal logic to time-indexed reality - **Uncertainty Reasoning**: Qualitative probability for everyday inference

20.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Fundamental eventuality framework and Rexist predicate
- **Chapter 8 (Logic Reified)**: Uses imply, not', and', or' for modal operations
- **Chapter 12 (Scales)**: Likelihood scales, Hi regions, ordering relationships
- **Chapter 16 (Time)**: atTime predicate and temporal anchoring
- **Future Psychology Chapters**: Modal framework for belief, planning, goal reasoning

20.6 Applications Mentioned:

- **Game Rules**: Tic-tac-toe possibility relative to game vs. physics constraints
- **Chess Necessity**: Required moves relative to chess rules vs. physical laws
- **Weather Prediction**: "Likely to rain" without precise probability
- **Everyday Uncertainty**: Qualitative likelihood judgments without mathematical basis
- **Temporal Examples**: Past likelihood (Booth acted alone), future likelihood (election outcomes)

20.7 Notable Design Decisions:

- **Constraint Relativity**: Possibility and necessity always relative to constraint sets
- **Qualitative Probability**: Partial ordering instead of numerical probabilities
- **Positive Modality Framework**: Unified treatment of inference-supporting modalities
- **Now as Anchor**: Connecting atemporal Rexist to temporal atTime
- **Non-Numerical Likelihood**: Avoiding precision that humans don't actually have
- **Scale-Based Hi Region**: "Likely" as high position on likelihood scale

20.8 Theoretical Significance:

Chapter 20 represents a sophisticated philosophical and logical foundation for reasoning about different modes of existence and uncertainty. The constraint-relative treatment of possibility and necessity provides a flexible framework that can accommodate different contexts (game rules, physical laws, social conventions) while maintaining logical rigor.

The qualitative approach to probability reflects how humans actually reason about uncertainty - with partially ordered judgments rather than precise numerical probabilities. The integration with scale theory enables contextual interpretation of likelihood claims while avoiding the false precision of numerical approaches.

The positive modality framework provides a unified foundation for different types of inference-supporting modalities, crucial for the cognitive theories that follow. The axiom schemas connecting different representational styles (unprimed/primed predicates, temporal arguments/atTime) enable flexible reasoning across different logical contexts.

This modal framework is essential for the psychological theories in Part III, providing the logical foundation for beliefs (which may not correspond to reality), goals (which represent desired possibilities), and plans (which reason about likelihood and necessity of outcomes). The chapter thus bridges pure logic and cognitive psychology by providing the modal concepts necessary for representing mental states and uncertain reasoning.

Chapter 21

Chapter 21: Knownledge management

- **112 axioms total** covering belief, knowledge, inference, justification, graded belief, assumptions, and mutual belief
- **13 main sections**: Objects of Belief, Belief, Belief Revision, Degrees of Belief, Assuming, Mind and Focus, Inference, Justification, Knowledge, Intelligence, Sentences/Domains, Expertise, Mutual Belief
- **All psychology** - first chapter in Part III focusing specifically on cognitive psychological phenomena

21.1 Key Features Identified:

1. **Foundational Belief Theory**:
   - Axioms 21.1-21.3: Concept-eventuality distinction and belief as relation to concepts
   - Axioms 21.4-21.8: Logic within belief contexts (conjunction, modus ponens, universal instantiation)
   - Axiom 21.9: Perception causes belief (defeasibly)
   - Axiom 21.10: Beliefs influence action through willing
2. **Belief Revision and Management**:
   - Axioms 21.19-21.27: Adding/deleting beliefs, recognizing inconsistencies, restoring consistency
   - Integration with AI belief revision literature and AGM postulates
   - Preference for minimal changes to belief sets when resolving contradictions
   - Agent abilities to manage their own knowledge
3. **Graded Belief Theory**:
   - Axioms 21.28-21.42: Degrees of belief using likelihood scales from Chapter 20
   - Graded belief operations for conjunction, disjunction, implication
   - Thresholds of belief determining bias toward belief/disbelief
   - Suspect, increaseBelief, and threshold-based absolute belief conversion

4. **Assumption Framework**:
   - Axioms 21.43-21.51: Assumptions as reasoning tool (hypothesis testing, accommodation)
   - Logic within assumption contexts (similar to belief but different causation)
   - Making/retracting assumptions as agent abilities
   - Assumptions leading to belief through consequence verification
5. **Mind Structure and Focus**:
   - Axioms 21.52-21.59: Mind with memory and focus of attention components
   - inm relation for mental containment, inFocus for attentional focus
   - thinkThat as conscious belief (belief in focus)
   - Foundation for attention-based cognitive processing
6. **Inference Theory**:
   - Axioms 21.60-21.72: Three modes of inference (deduction, abduction, induction)
   - Inference as causal relation from belief in premises to conscious belief in conclusion
   - Inference management (checking, suppressing, ignoring, contradictions, reaffirmation)
   - Confusion from inconsistent inferences
7. **Justification Taxonomy**:
   - Axioms 21.73-21.81: Multiple justification types (sound, partial, circular, poor, missing)
   - Justification as inference causing belief
   - Sound justification requires minimal proof and belief in all premises
   - Partial justification for fallible inference modes (abduction, induction)
8. **Knowledge as Justified True Belief**:
   - Axioms 21.84-21.94: Knowledge as true belief plus sound true justification
   - Learning, realizing, false positives/negatives
   - Intelligence scale based on inference abilities
   - Gettier problem addressed through justification truth requirement
9. **Sentences and Knowledge Domains**:
   - Axioms 21.95-21.105: Sentences with propositional content and claims
   - Knowledge domains characterized by predicate sets
   - Expertise scales and expert classification
   - Truth/falsity conditions for sentences vs. propositions
10. **Mutual Belief Framework**:
    - Axioms 21.106-21.113: Shared knowledge in communities
    - Mutual belief reflection property (believing that we mutually believe)
    - Copresence heuristic for establishing mutual belief
    - Extension to sentences and knowledge domains

21.2 Technical Sophistication:

- **Extensive Defeasibility**: 22 axioms use (etc) - more than any previous chapter
- **Concept-Eventuality Distinction**: Systematic treatment of mental representations

vs. world objects
- **Modal Integration**: Integration with likelihood scales and possibility theory from Chapter 20
- **Reified Cognitive States**: Extensive use of primed predicates for mental processes
- **Scale-Based Psychology**: Graded belief, intelligence, and expertise as scale positions

21.3 Complexity Distribution:

- Simple: 45 axioms (basic constraints, simple implications, defeasible rules)
- Moderate: 52 axioms (standard cognitive definitions, inference rules)
- Complex: 15 axioms (sophisticated epistemic definitions, multi-level cognitive processes)

21.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Cognitive Psychology**: Foundational concepts for all mental phenomena - **Artificial Intelligence**: Belief revision, inference, knowledge representation - **Philosophy of Mind**: Epistemic concepts, justification, knowledge conditions - **Social Cognition**: Mutual belief, shared knowledge, expertise attribution - **Natural Language Understanding**: Belief contexts, assumptions, communication

21.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Fundamental eventuality framework and Rexist
- **Chapter 7 (Substitution)**: Uses partialInstance for universal instantiation
- **Chapter 8 (Logic Reified)**: Uses and', not', imply' for belief operations
- **Chapter 12 (Scales)**: Graded belief, intelligence, and expertise scales
- **Chapter 15 (Causality)**: Causal relations in inference and belief formation
- **Chapter 19 (Persons)**: Uses perceive predicate and mind structure
- **Chapter 20 (Modality)**: Likelihood scales for graded belief theory

21.6 Applications Mentioned:

- **Belief Revision**: Discovering whales aren't fish, white dogs contradicting "all dogs are brown"
- **Inference Types**: Deductive logic, abductive explanation, inductive generalization
- **Graded Belief**: Mugger in next block scenario, threshold-based decisions
- **Assumptions**: Hypothesis testing, accommodation, social conventions (money)
- **Mutual Belief**: Copresence creating shared knowledge, cultural common ground

21.7 Notable Design Decisions:

- **Defeasible Logic**: Heavy emphasis on non-monotonic reasoning with (etc) conditions
- **Concept Mediation**: Belief relates agents to concepts, not directly to world eventualities
- **Anti-Logical Omniscience**: Explicit rejection of knowing all logical consequences
- **Threshold-Based Belief**: Graded belief converts to absolute belief via thresholds
- **Assumption-Belief Distinction**: Different causal patterns for assumptions vs. beliefs
- **Mind Structure**: Memory/focus distinction as foundation for attention
- **Justification Requirements**: Knowledge requires true justification, not just true belief

21.8 Theoretical Significance:

Chapter 21 represents the transition from abstract background theories to specifically psychological phenomena, establishing belief as the central cognitive relation. The extensive use of defeasible reasoning (22 axioms with etc) reflects the inherently non-monotonic nature of human cognition, where general rules have exceptions.

The concept-eventuality distinction provides a sophisticated solution to problems in philosophy of mind about mental representation, while the graded belief theory offers a qualitative alternative to precise probability assignments. The mutual belief framework enables modeling of shared knowledge and cultural common ground essential for communication and cooperation.

The integration with modal logic from Chapter 20 and scale theory from Chapter 12 demonstrates how abstract mathematical frameworks support concrete psychological modeling. The chapter's 112 axioms establish a comprehensive foundation for reasoning about human cognitive processes, setting the stage for more specific psychological theories in subsequent chapters.

This represents one of the most substantial formalizations of folk psychology concepts in the AI and cognitive science literature, providing both philosophical precision and computational tractability for modeling human-like reasoning and belief management.

Chapter 22

Chapter 22: Simility comparisons

22.1 Overview

- **28 axioms total** covering similarity, difference, analogies, pattern recognition, and cognitive comparison processes
- **3 main sections**: Similarity, Similarity of Structured Entities, Cognizing Similarities
- **Pure psychology** - sophisticated treatment of similarity as foundational cognitive operation

22.2 Key Features Identified:

22.2.1 1. Basic Similarity Theory:

- **Axioms 22.1-22.2**: `similarInThat` and `differentInThat` as fundamental relational concepts
- Properties held in common vs. properties that differ between entities
- Foundation for all higher-level similarity reasoning

22.2.2 2. Binary Similarity Framework (similar0):

- **Axioms 22.3-22.5**: Recursive co-definition of `similar0`, `simStr0`, `simPr0`
- Entities similar if they share properties OR have similar structure as eventualities
- Loop prevention through matched pairs tracking (parameter `m`)
- Binary judgment: entities are either similar or not

22.2.3 3. Graded Similarity Framework (similar1):

- **Axioms 22.6-22.11**: Enhanced similarity with explicit property set accumulation
- `similar1` tracks exactly which properties are shared (parameter `s`)
- More sophisticated than binary approach - enables quantitative similarity measurement

- **Axiom 22.11**: Equivalence theorem linking binary and graded approaches

22.2.4 4. Iterative Similarity Computation:

- **Axiom 22.8**: `iterArgs` - systematic traversal of eventuality arguments
- **Axiom 22.10**: `iterProps` - traversal of entity properties with inferential independence
- Complex bookkeeping to avoid infinite loops and ensure termination
- Builds up shared property sets incrementally

22.2.5 5. Similarity Scale Theory:

- **Axioms 22.12-22.17**: Integration with general scale framework from Chapter 12
- **Axiom 22.12**: Subset consistency - more shared properties = greater similarity
- **Axiom 22.14**: `similarityScale` based on `moreSimScale` partial ordering
- **Axioms 22.15-22.16**: `similar` and `different` as Hi/Lo regions on scale
- **Axiom 22.17**: Symmetry of similarity relation

22.2.6 6. Structured Entity Similarity:

- **Axiom 22.18**: `ceMapping` - composite entity mapping preserving structure
- **Axioms 22.19-22.20**: Pattern recognition through `exhibitPattern` and `commonPattern`
- Similarity based on shared structural organization, not just properties

22.2.7 7. Analogy Through Structure Mapping:

- **Axiom 22.21**: `cePredReplace` - systematic predicate transformation
- **Axiom 22.22**: `structureMapping` - composition of predicate replacement and entity mapping
- **Axiom 22.23**: `analogous` - entities with structure mappings between them
- Follows Gentner's (1983) structure mapping theory

22.2.8 8. Cognitive Similarity Processes:

- **Axiom 22.24**: `compare` - thinking about shared and differing properties
- **Axiom 22.25**: Comparison defeasibly causes similarity/difference judgments
- **Axiom 22.26**: Similarity scales as comparison metrics
- **Axioms 22.27-22.28**: `findPattern` and `drawAnalogy` as cognitive achievements

22.3 Technical Sophistication:

22.3.1 Recursive Complexity:

- Mutually recursive definitions with sophisticated loop prevention
- `simPr0`/`simStr0` and `simPr1`/`simStr1` co-define each other

- Parameter `m` tracks matched pairs to prevent infinite recursion

22.3.2 Graded vs. Binary Similarity:

- `similar0`: Existential approach (at least one shared property)
- `similar1`: Accumulative approach (explicit set of all shared properties)
- Enables both qualitative and quantitative similarity reasoning

22.3.3 Structural Sophistication:

- Beyond simple feature-based similarity to structural correspondence
- Handles complex compositional entities with components, properties, relations
- Systematic treatment of analogies through predicate mappings

22.3.4 Cognitive Integration:

- Bridges computational similarity algorithms with psychological processes
- Reified cognitive operations: `compare'`, `findPattern'`, `drawAnalogy'`
- Integration with belief and thinking frameworks from Chapter 21

22.4 Complexity Distribution:

- **Simple**: 3 axioms (basic definitions, symmetry)
- **Moderate**: 13 axioms (standard similarity definitions, cognitive processes)
- **Complex**: 12 axioms (recursive definitions, structural mappings, iterative procedures)

22.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Uses `eventuality`, `argn`, `pred` predicates extensively
- **Chapter 6 (Set Theory)**: Heavy use of set operations (`member`, `union`, etc.)
- **Chapter 7 (Substitution)**: Uses `subst` for property substitution in similarity
- **Chapter 8 (Logic Reified)**: Uses `Rexist` for property existence
- **Chapter 10 (Composite Entities)**: Pattern framework for structured similarity
- **Chapter 12 (Scales)**: Similarity scales using general scale theory
- **Chapter 15 (Causality)**: Causal relations in cognitive similarity processes
- **Chapter 21 (Knowledge Management)**: Uses `thinkOf`, `thinkThat'` predicates

22.6 Applications and Examples:

22.6.1 Feature-Based Similarity:

- Block comparison: A(red, square, large) vs B(red, round, large) vs C(red, square, small)

- A more similar to C than B (2 vs 1 shared properties)

22.6.2 Quantitative Measurements:

- Height similarity through shared threshold properties
- Pat(181cm) vs Chris(179cm) vs Kim(182cm) comparisons

22.6.3 Physics Problem Analysis:

- Ladder and man both exert forces with similar structure
- Recursive similarity through force→weight→objects→positions→endpoints

22.6.4 Analogy Examples:

- Planets orbiting Sun ↔ Students around professor
- Structure mapping preserves relational patterns while changing predicates

22.7 Notable Design Decisions:

22.7.1 Loop Prevention:

- Parameter `m` tracks already-matched pairs
- Essential for termination in recursive similarity checking
- Sophisticated bookkeeping throughout co-recursive definitions

22.7.2 Inferential Independence:

- `iterProps` ensures properties are inferentially independent
- Avoids counting "man"/"human"/"mammal"/"animal" as separate similarities
- Maintains meaningful similarity metrics

22.7.3 Symmetry vs. Asymmetry:

- Formal similarity relation is symmetric (Axiom 22.17)
- Acknowledges Tversky's asymmetry observations as discourse effects
- Distinguishes logical symmetry from pragmatic usage patterns

22.7.4 Graded Similarity:

- Parameter `s` accumulates shared properties for quantitative comparison
- Enables subset consistency (Axiom 22.12)
- Supports both binary and graded similarity judgments

22.8 Theoretical Significance:

Chapter 22 provides one of the most sophisticated formal treatments of similarity in cognitive science literature. The recursive, mutually-defined similarity predicates handle both simple feature-based similarity and complex structural correspondences.

The chapter's key innovation is the systematic treatment of similarity at multiple levels: 1. **Feature level**: Shared properties between entities 2. **Structural level**: Corresponding relationships and argument patterns 3. **Analogical level**: Systematic predicate mappings preserving structure 4. **Cognitive level**: Mental processes of comparison and pattern recognition

The graded similarity framework (`similar1`) advances beyond binary similarity judgments to enable quantitative similarity metrics while maintaining computational tractability. The integration with scale theory provides a principled foundation for similarity-based reasoning.

The treatment of analogies through structure mapping follows established cognitive science (Gentner 1983) while providing formal logical foundations. The systematic handling of composite entities enables similarity reasoning about complex structured objects.

The cognitive integration through reified comparison processes (`compare'`, `findPattern'`, `drawAnalogy'`) connects computational similarity algorithms with psychological theories of human similarity reasoning.

22.9 Technical Contributions:

22.9.1 Loop Prevention in Recursive Similarity:

- Novel formal treatment of the infinite regress problem in similarity
- Parameter `m` provides elegant solution to mutual recursion termination

22.9.2 Graded Similarity Accumulation:

- Parameter `s` enables explicit tracking of shared properties
- Supports both existential and universal similarity quantifications

22.9.3 Structural Similarity Framework:

- Systematic treatment of composite entity mappings
- Integration of pattern recognition with similarity theory

22.9.4 Scale-Based Similarity:

- Principled integration with general scale theory
- Enables context-dependent similarity thresholds through Hi/Lo regions

This chapter establishes similarity as a foundational cognitive operation supporting higher-level reasoning about patterns, analogies, and relationships. The formal framework provides both computational algorithms and psychological process models for human-like similarity reasoning.

Chapter 23

Chapter 23: Memory

38 axioms total covering memory storage, retrieval, accessibility, associations, multiple senses of remembering/forgetting, and psychological repression

6 main sections: Storing and Retrieving, Accessibility, Associations and Causing to Remember, The Meanings of "Remember" and "Forget", Remembering to Do, Repressing

All psychology domain - core cognitive infrastructure for memory operations and conscious thought

23.1 Key Features Identified:

23.1.1 1. Basic Memory Operations:

- **Axioms 23.1-23.5**: Foundational definitions of store/retrieve as state changes between focus and memory
- **Causality requirement**: Only way for concepts to enter memory is through storing (23.5)
- **Agency constraint**: Storage and retrieval are actions by agents (23.3-23.4)
- **State transition model**: Memory operations as changes between focus and memory states

23.1.2 2. Accessibility Framework:

- **Axioms 23.6-23.13**: Accessibility as partial ordering with agent-specific scales
- **Memory threshold**: Critical accessibility level below which retrieval is impossible
- **Functional relationship**: Each concept has exactly one accessibility value in a given memory
- **Scale integration**: Accessibility values exist on agent-specific scales

23.1.3 3. Importance-Accessibility Connection:

- **Axiom 23.14**: Defeasible monotonic relation between concept importance and accessibility
- **Goal-driven memory**: More important concepts (relative to goals) are more accessible

- **Emotional significance**: Explains why forgetting anniversaries causes anger
- **Cross-chapter integration**: Connects memory theory to goal theory (Chapter 28)

23.1.4 4. Association Theory:

- **Axioms 23.15-23.21**: Comprehensive framework for conceptual associations
- **Causal associations**: Causes/effects are associated (23.18)
- **Logical associations**: Implication antecedents/consequents are associated (23.19)

- **Structural associations**: Eventualities associate with their arguments (23.20)
- **Accessibility propagation**: Focus on one concept raises accessibility of associated concepts (23.21)

23.1.5 5. Hierarchical Remembering:

- **Axioms 23.22-23.30**: Four distinct senses of "remember" with entailment relations
- **Remember1**: Accessible but not necessarily in focus
- **Remember2**: Actually retrieved from memory
- **Remember3**: Agent causes the retrieval
- **Remember4**: Agent succeeds in deliberate retrieval attempt
- **Hierarchical entailment**: Each sense implies all previous senses

23.1.6 6. Forgetting Taxonomy:

- **Axioms 23.31-23.34**: Three distinct senses of "forget"
- **Forget1**: Accessibility drops below memory threshold
- **Forget2**: Failure to retrieve when beneficial
- **Forget3**: Intentional removal from focus
- **Complementarity**: Forget1 negates all senses of remembering (23.32)

23.1.7 7. Temporal Memory (Remembering to Do):

- **Axioms 23.35-23.36**: Framework for planned action execution
- **Focus requirement**: Actions require being in focus at execution time
- **Plan integration**: Remembering to do becomes part of goal-directed planning
- **Temporal coordination**: Retrieval must cause focus at the right time

23.1.8 8. Repression Theory:

- **Axioms 23.37-23.38**: Freudian-inspired model of memory repression
- **Accessibility blocking**: Repressed concepts are in memory but not accessible
- **Emotional causation**: Repression is defeasibly caused by concepts causing unhappiness
- **Unconscious processes**: Framework for non-accessible but stored memories

23.2 Technical Sophistication:

- **State-based modeling**: Memory operations as transitions between cognitive states
- **Scale theory integration**: Accessibility and importance as scale positions
- **Reified cognitive processes**: Extensive use of primed predicates for mental operations
- **Defeasible reasoning**: 4 axioms use (etc) for non-monotonic memory principles
- **Lexical grounding**: Systematic mapping between theory predicates and natural language

23.3 Complexity Distribution:

- **Simple**: 12 axioms (basic constraints, type requirements, implications)
- **Moderate**: 14 axioms (standard definitions with existential quantification)
- **Complex**: 12 axioms (multi-layered state transitions, causal chains)

23.4 Conceptual Importance:

This chapter provides essential cognitive infrastructure for: - **Artificial Intelligence**: Memory architectures, retrieval mechanisms, attention models - **Cognitive Psychology**: Formal models of human memory processes and accessibility - **Natural Language Processing**: Grounding for memory-related linguistic expressions - **Planning Systems**: Temporal memory for action execution and goal achievement - **Psychotherapy**: Formal framework for understanding repression and memory disorders

23.5 Cross-Chapter Connections:

- **Chapter 21 (Knowledge Management)**: Uses mind structure with memory and focus components
- **Chapter 28 (Goals)**: Importance scales and goodFor predicate for memory prioritization

- **Chapter 24 (Envisioning)**: thinkOf predicate for conscious thought processes
- **Chapter 49 (Emotions)**: unhappy predicate for emotional causation in repression
- **Chapter 12 (Scales)**: Scale theory foundation for accessibility and importance
- **Chapter 15 (Causality)**: Causal relations in memory operations and associations

23.6 Applications Mentioned:

- **Memory strategies**: Alphabet method for name recall, physical reminders for actions
- **Emotional memory**: Anniversary forgetting causing relationship conflict
- **Association chains**: Conjunctions/disjunctions triggering related concept recall
- **Repressed memories**: Traumatic experiences blocked from conscious access

- **Planned actions**: Grocery shopping as part of dinner preparation goal

23.7 Notable Design Decisions:

- **No Platonic ideals**: Explicit rejection of pre-existing memory contents (23.5)
- **Agent-relative accessibility**: No cross-agent accessibility comparisons
- **Multiple forgetting senses**: Distinguishing inability, failure, and intentional forgetting
- **Defeasible associations**: Association symmetry with exceptions via (etc)
- **Reified mental states**: Primed predicates for all memory operations
- **Temporal action coupling**: Required focus for action execution at specific times

23.8 Theoretical Significance:

Chapter 23 establishes memory as the foundational cognitive mechanism connecting perception, knowledge, goals, and action. The accessibility framework provides a psychologically realistic model of memory limitations, while the association theory explains memory retrieval through conceptual connections.

The hierarchical analysis of "remember" and "forget" demonstrates the sophisticated relationship between formal psychological theories and natural language semantics. The multiple senses capture genuine psychological distinctions while maintaining logical coherence through entailment relations.

The integration of importance-based accessibility provides a goal-driven model of memory prioritization, connecting cognitive psychology with practical artificial intelligence applications. The repression framework addresses unconscious mental processes while maintaining computational tractability.

This represents one of the most comprehensive formal treatments of human memory processes, bridging cognitive science, artificial intelligence, and natural language understanding within a unified logical framework.

Chapter 24

Chapter 24: Envisioning

24.1 Overview

- **52 axioms total** covering thinking, prediction, explanation, causal systems, envisioned causal systems, and belief interactions
- **6 main sections**: Thinking Of, Causal Systems, Contiguous Causal Systems, Envisioned Causal Systems, Envisionment and Belief, Other Varieties of Thinking
- **Mixed domains** - 16 background theory axioms (causal systems formalization) + 36 psychology axioms (cognitive processes)

24.2 Key Features Identified:

24.2.1 1. Foundational Thinking Theory:

- **ThinkOf Definition (24.1)**: Core cognitive relation - having concepts in focus of attention
- **Perception-Thought Link (24.2-24.3)**: Perception defeasibly causes conscious thought and belief in existence
- **Association Chains (24.4)**: Associated entities trigger thinking cascades through defeasible causation
- **Truth Independence**: Thinking is independent of truth, belief, and likelihood - agents can think of anything

24.2.2 2. Prediction and Explanation Framework:

- **Prediction Theory (24.5-24.8)**: Formal temporal logic - believing something will occur at future times, with validation/invalidation
- **Causal Prediction (24.6)**: Believing cause + thinking cause obtains → defeasibly predict effect
- **Explanation Definition (24.9)**: Explaining e with e1 means thinking e1 caused e

- **Alternative Reasoning (24.10-24.11)**: Multiple causes/effects trigger disjunctive thinking about OR combinations

24.2.3 3. Sophisticated Causal Systems Theory:

- **Graph-Theoretic Foundation**: Causal systems as directed AND-OR graphs with eventualities as nodes
- **Causal System Arcs (24.15)**: Either causallyInvolved relations or disjunctive branch relations
- **Branch Structure (24.12-24.14)**: Formal disjunctive branches with reified disjunct relations
- **System Properties**: Connected, branchless, isolated node definitions with graph-theoretic precision

24.2.4 4. Contiguous Causal System Operations:

- **OneArcDiff (24.25)**: Adding/removing single causal arcs for incremental system modification
- **ResolveBranch (24.27)**: Complex operation for resolving disjunctive branches to single alternatives
- **Contiguity Definition (24.28)**: Two systems are contiguous if related by oneArcDiff or resolveBranch
- **Incremental Thinking**: Framework for how agents move between related causal understanding states

24.2.5 5. Envisioned Causal System Sequences:

- **ECS Definition (24.30)**: Agent thinks of all eventualities, believes or focuses on all relations
- **Background Beliefs (24.31)**: Constraining envisionments with unchallengeable background assumptions
- **ECS Sequences (24.32)**: Temporal sequences of contiguous causal systems with change-of-state links
- **Directional Envisioning (24.34-24.36)**: EnvisionFromTo, envisionFrom, envisionTo for goal-directed thinking

24.2.6 6. Belief-Envisionment Integration:

- **System Belief (24.39)**: Believing causal system = believing conjunction of eventualities + relations
- **Graded Belief Dynamics (24.42-24.46)**: Element belief changes cause proportional system belief changes
- **Verification Effects (24.47-24.50)**: Finding believed causes/effects or falsifying predictions updates system beliefs

- **Current World Understanding (24.51-24.52)**: ECS containing all perceived eventualities = agent's cwu

24.3 Technical Sophistication:

24.3.1 Extensive Reification:

- 16 primed predicates for cognitive and causal processes: `thinkOf'`, `thinkThat'`, `perceive'`, `cause'`, `predict'`, `ecs'`, `causallyInvolved'`, `disjunct'`, `csArc'`, `changeGBel'`, `change'`, `not'`, `gbel'`, and others
- Enables sophisticated temporal and causal reasoning about mental processes

24.3.2 Graph-Theoretic Precision:

- Formal AND-OR graphs with precise connectivity, branch resolution, and incremental modification operations
- Most mathematically sophisticated psychological theory in the corpus so far

24.3.3 Defeasible Psychology:

- 12 axioms use `(etc)` conditions reflecting non-monotonic nature of thinking, association, prediction, and belief dynamics
- Captures inherent uncertainty in psychological processes

24.3.4 Process Integration:

- Links perception → thinking → prediction/explanation → belief revision in coherent causal framework
- Sophisticated integration across multiple cognitive domains

24.3.5 Recursive Definitions:

- CausallyLinked uses recursive graph traversal for connectivity determination
- Complex resolveBranch operation with multiple conditional branches

24.4 Complexity Distribution:

- **Simple: 12 axioms** (basic definitions, simple existence claims, symmetry properties)
- **Moderate: 23 axioms** (standard cognitive and system definitions with moderate complexity)
- **Complex: 17 axioms** (sophisticated graph operations, recursive definitions, multi-conditional processes)

24.5 Conceptual Importance:

24.5.1 Cognitive Architecture:

Provides the most detailed formal model of causal thinking in cognitive science literature. Shows how agents construct, modify, and reason with mental models of causal relationships in systematic, incremental ways.

24.5.2 AI Reasoning Systems:

Offers implementable framework for causal reasoning, explanation generation, and belief revision. The graph-theoretic foundation enables efficient algorithms for causal inference and mental model construction.

24.5.3 Philosophy of Mind:

Formalizes key intuitions about mental representation, causal reasoning, and the relationship between thinking and believing. Addresses fundamental questions about how minds represent and reason about causality.

24.5.4 Computational Modeling:

Provides mathematical framework for implementing human-like causal reasoning in AI systems, with precise operations for mental model construction, modification, and evaluation.

24.6 Cross-Chapter Connections:

- **Chapter 15 (Causality)**: Uses `causallyInvolved` and `cause` predicates as foundation
- **Chapter 21 (Belief Management)**: Uses `believe`, `focusOfAttention`, `gbel` for cognitive integration
- **Chapter 23 (Accessibility)**: Uses `associated` predicate for thinking chain triggers
- **Chapters 25+ (Future)**: Provides foundation for explanation, planning, and other cognitive processes

24.7 Applications Mentioned:

- **Causal Reasoning**: Medical diagnosis, scientific hypothesis formation, everyday explanation
- **Mental Simulation**: What-if reasoning, counterfactual thinking, planning scenarios
- **Belief Revision**: Theory change in science, learning from experience, diagnostic reasoning
- **Mathematical Thinking**: Tracing implicational networks, theorem proving, logical reasoning

24.8 Notable Design Decisions:

24.8.1 Thinking-Belief Independence:

Explicit separation allows agents to think about false, unlikely, or disbelieved propositions - crucial for counterfactual reasoning and imagination.

24.8.2 Graph-Theoretic Rigor:

Mathematical precision enables computational implementation while maintaining psychological plausibility through incremental operations.

24.8.3 Defeasible Associations:

Thinking chains follow associations defeasibly, capturing both systematic and creative aspects of human thought.

24.8.4 Temporal Anchoring:

Prediction and validation tied to specific times, enabling precise treatment of temporal reasoning and belief dynamics.

24.8.5 Incremental Modification:

OneArcDiff and resolveBranch operations capture how humans modify causal understanding gradually rather than rebuilding entire mental models.

24.8.6 Background Belief Constraints:

Allows for context-dependent reasoning where some beliefs remain fixed while others are explored or revised.

24.9 Theoretical Significance:

Chapter 24 represents a major theoretical contribution to formal cognitive science, providing the first mathematically rigorous account of causal thinking that integrates:

1. **Mental Representation**: How causal knowledge is structured as graph-like mental models
2. **Cognitive Processes**: How thinking, prediction, and explanation operate over these representations

3. **Belief Dynamics**: How experience and reasoning modify confidence in causal models
4. **Incremental Reasoning**: How agents modify understanding through small, systematic changes

The extensive use of reified predicates (16 different primed predicates) reflects the process-oriented nature of causal thinking - it's not just about static causal beliefs but about the dynamic mental processes that construct, evaluate, and modify causal understanding.

The graph-theoretic foundation provides both psychological realism (through incremental operations) and computational tractability (through well-defined algorithms for graph traversal and modification). The defeasible framework captures the inherently non-monotonic nature of causal reasoning where new evidence can overturn previous conclusions.

The integration of perception, thinking, prediction, explanation, and belief revision into a unified framework represents a major advance in understanding how human causal cognition operates as an integrated system rather than separate modules.

This chapter establishes causal thinking as the central cognitive process for understanding and predicting the world, providing the foundation for more specialized cognitive abilities like planning, explanation, and problem-solving covered in subsequent chapters.

24.10 Pattern Analysis:

- **35 Definitions** - Comprehensive formalization of causal thinking concepts
- **12 Defeasible Rules** - Non-monotonic aspects of thinking and belief dynamics
- **3 Existence Claims** - Basic facts about agent capabilities
- **2 Argument Structure** - Symmetry and structural properties
- **1 Recursive Definition** - Graph connectivity with recursive traversal

The chapter represents the most mathematically sophisticated treatment of psychological processes in the corpus, bridging formal graph theory with cognitive psychology to provide both theoretical insight and computational implementability for human-like causal reasoning systems.

Chapter 25

Chapter 25: Explanation

25.1 Overview

- **22 axioms total** covering explanations, mysteries, explanation processes, and explanation failures
- **3 main sections**: Explanations and Mysteries, The Explanation Process, Explanation Failures
- **Pure psychology domain** - focuses on cognitive processes of causal reasoning and explanation generation

25.2 Key Features Identified:

25.2.1 1. Foundational Explanation Theory:

- **Axioms 25.1-25.4**: Basic existence claims - agents can explain some things, people do explain things, mysteries exist for everyone
- **Mystery Definition (25.3)**: Formal definition linking inability to explain with agent's current world understanding
- **Explanation as Causal Belief**: Core insight that explaining e with e1 means believing e1 caused e

25.2.2 2. Multiple Explanations and Preferences:

- **Axiom 25.6**: Formal recognition that eventualities can have multiple possible explanations
- **Better Explanations (25.7)**: Graded belief determines explanation quality - higher belief = better explanation

- **Domain Preferences (25.8)**: Agents prefer explanations from certain knowledge domains (theological vs. biological vocabularies)

- **Best Explanation Framework (25.9-25.10)**: Partial ordering leads to best explanation adoption through defeasible causation

25.2.3 3. Idealized Explanation Process Model:

- **Process Breakdown (25.11)**: Four-stage temporal process: adopt goal → generate candidates → assess candidates → adopt explanation
- **Goal Adoption (25.12-25.13)**: Triggered by unpredicted events that require causal understanding
- **Generation Phase (25.14)**: Think of possible causes given current world understanding
- **Assessment Phase (25.15)**: Compare explanations pairwise to determine relative quality
- **Adoption/Rejection (25.16-25.18)**: Adopt best explanation, defeasibly reject alternatives

25.2.4 4. Systematic Failure Analysis:

- **General Failure Definition (25.19)**: Failure occurs when all explanation attempts fail
- **Three Failure Points**:
  - **Generation Failure (25.20)**: Cannot think of any candidate explanations
  - **Assessment Failure (25.21)**: Generate candidates but fail to evaluate them
  - **Adoption Failure (25.22)**: Evaluate candidates but fail to commit to any explanation

25.3 Technical Sophistication:

25.3.1 Extensive Reification:

- 11 primed predicates for cognitive processes: `explain'`, `adoptGoalToExplain'`, `generateExplanations'`, `assessExplanations'`, `adoptExplanation'`, `rejectExplanation'`, `bestExplanationFor'`, `cause'`, `believe'`, `goal'`, `changeTo'`, `not'`
- Enables temporal reasoning about explanation processes using `before` predicate

25.3.2 Defeasible Psychology:

- 6 axioms use `(etc)` conditions reflecting non-monotonic nature of explanation behavior
- Covers agent goals, explanation preferences, process idealization, and alternative rejection

25.3.3 Process Integration:

- Links to **Chapter 21** (knowledge domains, graded belief), **Chapter 24** (explain definition), **Chapter 28** (goals, trying, failure)
- Sophisticated integration of belief revision, preference reasoning, and goal management

25.3.4 Preference Mechanisms:

- **Quantitative**: Graded belief degrees determine explanation quality

- **Qualitative**: Knowledge domain membership creates systematic preferences
- **Agent-relative**: `betterExplanationFor` includes agent parameter recognizing subjective variation

25.4 Complexity Distribution:

- **Simple: 6 axioms** (basic existence claims, simple failure conditions)
- **Moderate: 8 axioms** (standard cognitive definitions, process steps)
- **Complex: 8 axioms** (sophisticated preference mechanisms, recursive process definitions)

25.5 Conceptual Importance:

25.5.1 Cognitive Architecture:

Provides detailed process model for one of the most important cognitive abilities - causal reasoning and explanation generation. Shows how agents move from observation through hypothesis generation to belief formation.

25.5.2 AI Reasoning Systems:

Offers formal framework for automated explanation systems, diagnostic reasoning, and abductive inference. The failure analysis provides debugging framework for explanation systems.

25.5.3 Philosophy of Science:

Formalizes intuitions about explanation competition, theory preference, and the role of background knowledge domains in scientific reasoning.

25.5.4 Natural Language Understanding:

Explains how humans generate and evaluate explanatory discourse, critical for systems that need to produce or understand explanations in natural language.

25.6 Cross-Chapter Connections:

- **Chapter 21 (Belief Management)**: Uses `gbel` for graded belief, `knowledgeDomain` for preferences
- **Chapter 24 (Envisioning)**: Builds on basic `explain` predicate definition
- **Chapter 28 (Goals)**: Uses `goal`, `try`, `fail` predicates for process control
- **Chapter 15 (Causality)**: Explanation as belief in causal relations
- **Chapter 19 (Persons)**: Agent abilities and current world understanding

25.7 Applications Mentioned:

- **Unpredicted Events**: Learning something unexpected triggers explanation goals
- **Multiple Theories**: Same phenomenon can have competing explanations (natural selection vs. intelligent design)
- **Domain Preferences**: Theological vs. biological vocabularies for life phenomena
- **Diagnostic Reasoning**: Medical diagnosis as explanation generation and assessment

25.8 Notable Design Decisions:

25.8.1 Process Idealization:

Explicitly acknowledges that the four-stage model is idealized - real explanation processes may be messier, but the model captures the essential structure.

25.8.2 Defeasible Alternative Rejection:

Adopting one explanation defeasibly (not necessarily) causes rejection of alternatives - allows for agents who maintain multiple competing hypotheses.

25.8.3 Failure Point Analysis:

Systematic analysis of where explanation processes can break down, enabling diagnostic reasoning about reasoning failures.

25.8.4 Agent-Relative Quality:

Explanation quality is always relative to an agent - no objective "best explanation" independent of agent beliefs and preferences.

25.8.5 Temporal Ordering:

Uses `before` predicate to capture essential temporal constraints in explanation processes - cannot assess before generating, cannot adopt before assessing.

25.9 Theoretical Significance:

Chapter 25 provides one of the most detailed formal models of explanation processes in the cognitive science literature. It bridges philosophical work on explanation (inference to the best explanation, explanatory virtues) with computational AI work on abductive reasoning and diagnostic systems.

The integration of quantitative factors (graded belief) with qualitative factors (knowledge domain preferences) offers a sophisticated account of how explanation quality is determined. The systematic

failure analysis provides both theoretical insight into explanation breakdowns and practical guidance for building robust explanation systems.

The heavy use of reified predicates (11 out of 22 axioms) reflects the process-oriented nature of explanation - it's not just about static explanation relations but about the dynamic cognitive processes that generate, evaluate, and adopt explanations.

The defeasible framework captures the inherently non-monotonic nature of explanation reasoning - new evidence can overturn explanations, multiple explanations can coexist, and explanation preferences can vary across agents and contexts.

This chapter establishes explanation as a central cognitive process requiring sophisticated integration of belief management, goal reasoning, causal understanding, and preference mechanisms. It provides the foundation for understanding how agents make sense of their world through causal reasoning and hypothesis formation.

25.10 Pattern Analysis:

- **11 Definitions** - Core predicates for explanation processes
- **6 Defeasible Rules** - Non-monotonic aspects of explanation behavior

- **3 Existence Claims** - Basic facts about agent explanation abilities
- **2 Goal Reasoning** - Integration with goal management framework

The chapter represents a mature integration of philosophical insights about explanation with computational process models, providing both theoretical precision and implementational guidance for explanation-capable AI systems.

Chapter 26

Chapter 26: Managing Expectations

26.1 Overview

- **9 axioms total** covering expectations, justifications, expectation management, and confirmation/violation
- **1 main section** with focused treatment of expectation theory
- **Pure psychology domain** - all axioms deal with cognitive processes and temporal beliefs

26.2 Key Features Identified:

26.2.1 1. Foundational Expectation Theory:

- **Expect Definition (26.1)**: Expectation as temporally anchored belief - believing at t0 that e will occur at future time t1
- **Temporal Constraint**: Essential `before t0 t1` constraint ensuring expectations are future-directed
- **Belief Constitution**: Expectation constituted by belief e2 that e will occur at t1, with causal generation relation
- **Reified Process**: `expect'` allows temporal reasoning about expectation formation and management

26.2.2 2. Complementary Unexpectation:

- **Unexpect Definition (26.2)**: Failure to believe event will occur at future time - not merely absence of expectation
- **Temporal Specificity**: Like expectation, unexpectation is temporally anchored to specific future times
- **Cognitive Reality**: Unexpectation as positive cognitive state, not just lack of expectation

26.2.3 3. Sophisticated Justification Framework:

- **Justification Theory (26.3)**: Expectations justified by envisioned causal systems (ecs) that support them
- **Causal Integration**: Current world understanding (cwu) at t0 causally generates expectation when it contains expected eventuality
- **ECS Foundation**: Links expectation management to sophisticated causal reasoning from Chapter 24
- **Membership Requirement**: Expected eventuality must be member of eventualities in justifying causal system

26.2.4 4. Dynamic Expectation Management:

- **Add Expectations (26.4)**: Extending envisioned causal systems to point where target eventuality is predicted
- **Remove Expectations (26.5)**: Inverse operation - moving from justified to unjustified expectation state
- **Change-of-State Framework**: Both operations involve `change` between different ecs states
- **Temporal Flexibility**: Operations can target different future times and modify different aspects of causal understanding

26.2.5 5. Agent Capabilities:

- **Add Ability (26.6)**: People have ability to add expectations subject to constraint conditions
- **Remove Ability (26.7)**: People have ability to remove expectations subject to constraint conditions
- **Individual Differences**: Text notes expertise effects - some better at stock market/election prediction
- **Skill Variation**: Better general prediction ability from projecting further forward and considering relevant factors

26.2.6 6. Expectation Outcomes:

- **Confirmation (26.8)**: When expected time arrives, expectation confirmed if event actually occurs
- **Violation (26.9)**: When expected time arrives, expectation violated if event does not occur
- **Surprise Distinction**: Expectation violation differs from out-of-the-blue surprises - requires prior specific expectation

26.3 Technical Sophistication:

26.3.1 Temporal Logic Integration:

- 7 different temporal predicates: `atTime`, `atTime'`, `before`, temporal variables t0/t1/t2
- Precise temporal anchoring of beliefs, expectations, and causal reasoning
- Future-directed constraint ensuring expectations properly oriented toward future

26.3.2 Extensive Reification:

- 7 primed predicates for temporal and cognitive processes: `expect'`, `atTime'`, `believe'`, `cwu0'`, `ecs'`, `addExpectation'`, `removeExpectation'`
- Enables sophisticated reasoning about expectation processes and their temporal properties

26.3.3 Causal System Integration:

- Deep integration with Chapter 24 envisioning framework through `ecs'`, `cwu0'`, `eventualitiesOf`
- Expectations grounded in agent's causal understanding of world rather than isolated beliefs
- Change operations maintain coherence with causal reasoning processes

26.3.4 No Defeasible Reasoning:

- Unique among psychology chapters - no axioms use `(etc)` conditions
- Focuses on precise temporal and belief relationships rather than defeasible patterns
- Reflects mathematical precision appropriate for temporal logic

26.4 Complexity Distribution:

- **Simple: 4 axioms** (basic agent abilities, confirmation/violation outcomes)
- **Moderate: 1 axiom** (unexpectation definition)
- **Complex: 4 axioms** (core expectation definition, justification theory, expectation management operations)

26.5 Conceptual Importance:

26.5.1 Cognitive Architecture:

Provides precise temporal framework for future-directed cognition. Expectations emerge from causal understanding and can be systematically managed through ecs modifications. Links anticipation to sophisticated mental model reasoning.

26.5.2 AI Planning Systems:

Formal framework for expectation-based reasoning in autonomous agents. Enables systematic belief revision when expectations violated, and provides foundation for learning from predictive failures.

26.5.3 Philosophy of Mind:

Addresses fundamental questions about temporal cognition and future-directed mental states. Clarifies relationship between believing, expecting, and temporal reasoning about possibilities.

26.5.4 Predictive Processing:

Aligns with modern cognitive science emphasis on predictive minds. Formal account of how expectations are generated from internal models and updated based on experience.

26.6 Cross-Chapter Connections:

- **Chapter 24 (Envisioning)**: Uses `ecs`, `cwu0`, `change`, `eventualitiesOf` as foundation for expectation justification
- **Chapter 21 (Belief Management)**: Uses `believe'` for temporal belief states
- **Temporal Framework**: Integrates with temporal logic infrastructure from earlier chapters
- **Future Chapters**: Likely foundation for planning, goal management, and action selection

26.7 Applications Mentioned:

- **Domain Expertise**: Stock market prediction, election forecasting based on domain knowledge and probabilities
- **Individual Differences**: Variation in prediction abilities based on forward projection skills and factor consideration
- **Surprise vs. Violation**: Walking tackled by stranger (surprise) vs. predicted outcome not occurring (violation)
- **Causal Prediction**: Using causal system understanding to generate specific temporal expectations

26.8 Notable Design Decisions:

26.8.1 Temporal Anchoring:

Expectations always anchored to specific times (t0 for formation, t1 for occurrence) enabling precise temporal reasoning and outcome evaluation.

26.8.2 Belief Constitution:

Expectations constituted by beliefs rather than separate cognitive category - maintains theoretical parsimony while allowing temporal specialization.

26.8.3 Causal Grounding:

Expectations justified by envisioned causal systems rather than isolated predictions - integrates with sophisticated causal reasoning framework.

26.8.4 Process Reification:

Expectation formation, management, and outcomes all reified as eventualities enabling temporal and causal reasoning about expectation processes themselves.

26.8.5 Individual Differences:

Explicit recognition that expectation abilities vary across agents based on expertise, experience, and cognitive skills.

26.8.6 Surprise Distinction:

Careful distinction between expectation violations (predicted A, got B) and surprises (no specific expectation, unexpected event) maintains theoretical precision.

26.9 Theoretical Significance:

Chapter 26 provides a mathematically precise account of expectation that integrates temporal logic, belief theory, and causal reasoning into a coherent framework for future-directed cognition.

The chapter's strength lies in its integration with the sophisticated causal reasoning framework from Chapter 24. Rather than treating expectations as isolated predictions, they emerge from and are justified by agent's envisioned causal systems. This provides both psychological realism (expectations based on causal understanding) and computational tractability (systematic ecs modification operations).

The temporal precision enables exact specification of when expectations form, when they target, and when they are confirmed or violated. This mathematical rigor supports both theoretical analysis and practical implementation in AI systems requiring temporal reasoning and predictive capabilities.

The reification of expectation processes (through 7 primed predicates) allows for sophisticated reasoning about expectation management itself - agents can reason about their own expectation formation, modification, and evaluation processes. This meta-cognitive capability is essential for learning and belief revision.

The integration of individual differences (expertise, skill variation) acknowledges that expectation abilities are not uniform across agents while maintaining the formal framework's precision. This enables modeling of learning, expertise development, and skill-based prediction differences.

The careful distinction between expectation violations and surprises reflects deep understanding of cognitive phenomenology - the different experiential and computational qualities of predicted-but-wrong vs. unpredicted events.

This chapter establishes expectations as temporally precise, causally grounded, and systematically manageable mental states that bridge current understanding with future-directed action and belief revision. It provides the foundation for planning, goal management, and adaptive behavior in uncertain environments.

26.10 Pattern Analysis:

- **7 Definitions** - Core framework for expectation theory
- **2 Existence Claims** - Basic agent capabilities for expectation management
- **0 Defeasible Rules** - Unique precision focus without defeasible patterns

The chapter represents a focused, mathematically rigorous treatment of temporal cognition that maintains deep integration with the broader causal reasoning framework while providing precise tools for modeling future-directed mental states and their management.

Chapter 27

Chapter 27: Other-Agent Reasoning

27.1 Overview

- **8 axioms total** covering theory of mind, introspection, other-agent reasoning, and mental state models
- **1 main section** with focused treatment of mental state reasoning
- **Pure psychology domain** - all axioms deal with cognitive processes about minds

27.2 Key Features Identified:

27.2.1 1. Foundational Theory of Mind Framework:

- **EnvisionMentalState (27.1)**: Core definition - agent a envisions agent b's mental state e through causal system containing b's thinking
- **Causal System Integration**: Mental state reasoning built on established envisioned causal system (ecs) infrastructure from Chapter 24
- **Mental Events as Events**: Mental events treated as particular kind of event within general causal framework - no special ontological category
- **Systematic Approach**: Uses existing causal reasoning machinery rather than separate theory of mind module

27.2.2 2. Agent Capabilities and Limitations:

- **General Ability (27.2)**: People defeasibly able to envision other people's mental states, subject to constraint conditions
- **Individual Variation**: Ability depends on implicit constraints c - captures individual differences in theory of mind skills
- **Failure Modes (27.4)**: Explicit recognition that people sometimes fail at other-agent reasoning

- **Constraint Sensitivity**: Abilities are context-dependent and can be limited by various factors

27.2.3 3. Other-Agent vs. Self Distinction:

- **Other-Agent Definition (27.3)**: Simple constraint that reasoning agent and target agent must be different (nequal a b)
- **Introspection (27.5)**: Special case where agent envisions own mental states (a = a)
- **Cognaesthetic Sense**: People have partial observation/perception of their own thought processes
- **Introspection Failure (27.6)**: Even self-knowledge can fail - agents not always transparent to themselves

27.2.4 4. Mental Models and Stereotypes:

- **General Mental Models (27.7)**: People have beliefs in causal systems involving how other people think
- **Group Models (27.8)**: More specific models for how members of certain groups think
- **Belief-Based**: Mental models are beliefs in causal systems rather than direct access to others' minds
- **Stereotyping Foundation**: Formal foundation for understanding social cognition and group-based reasoning

27.2.5 5. Cognitive Advice Framework:

- **Advice Reception**: Brief mention that people can take cognitive advice - adopting beliefs in causal systems containing mental acts by the agent
- **Self-Modification**: Agents can modify their own mental processes based on external recommendations
- **Meta-Cognitive Control**: Links to broader framework of cognitive self-management

27.3 Technical Sophistication:

27.3.1 Minimal Reification:

- Only 4 primed predicates used: `thinkOf'`, `envisionMentalState'`, `otherAgentReason'`, `introspect'`
- Simpler than most psychology chapters, reflecting straightforward application of existing framework

27.3.2 Causal System Foundation:

- Builds directly on Chapter 24 infrastructure: `ecs`, `eventualitiesOf`, `member`

- Mental state reasoning as specialized application of general causal reasoning
- No separate theoretical machinery needed for theory of mind

27.3.3 Minimal Defeasible Reasoning:

- Only 1 axiom uses `(etc)` conditions (27.2) - mostly precise definitional framework
- Reflects that theory of mind concepts have clear structural definitions

27.3.4 Individual Differences:

- Constraint parameter c in ability axiom captures variation in theory of mind skills
- Formal recognition that not all agents equally capable of mental state reasoning

27.4 Complexity Distribution:

- **Simple: 4 axioms** (basic definitions, introspection, failure modes)
- **Moderate: 3 axioms** (mental state envisioning, abilities, general mental models)
- **Complex: 1 axiom** (group mental models with nested quantification)

27.5 Conceptual Importance:

27.5.1 Theory of Mind:

Provides formal account of fundamental social cognitive ability. Shows how agents can reason about unobservable mental states of others using causal system envisioning rather than direct mental access.

27.5.2 Social Cognition:

Foundation for understanding stereotyping, empathy, perspective-taking, and social interaction. Mental models of groups provide formal basis for social categorization and expectation formation.

27.5.3 Introspection:

Treats self-knowledge as special case of mental state reasoning rather than privileged access. Explains why introspection can fail and why self-knowledge is sometimes difficult.

27.5.4 AI Social Reasoning:

Provides implementable framework for artificial agents to reason about human mental states and develop appropriate social responses based on mental state attribution.

27.6 Cross-Chapter Connections:

- **Chapter 24 (Envisioning)**: Uses `ecs`, `eventualitiesOf`, `member` as foundation for mental state reasoning
- **Chapter 21 (Belief Management)**: Uses `believe` for mental model representation
- **Chapter 28 (Goals)**: References `fail` predicate for failure modes
- **Future chapters**: Likely foundation for social planning, communication, and cooperative behavior

27.7 Applications Mentioned:

- **Cognaesthetic Sense**: Partial observation of one's own thought processes - foundation for metacognition
- **Cognitive Advice**: Taking recommendations about how to think or reason from others
- **Group Stereotypes**: Mental models of how members of specific groups think and behave
- **Social Interaction**: Understanding others' mental states for successful communication and cooperation

27.8 Notable Design Decisions:

27.8.1 Mental Events as Events:

No special ontological category for mental events - they're just particular kinds of events in the general causal framework. Maintains theoretical parsimony.

27.8.2 Causal System Foundation:

Mental state reasoning built on established causal reasoning infrastructure rather than separate module. Enables integration with broader cognitive architecture.

27.8.3 Failure Modes:

Explicit treatment of both other-agent reasoning failure and introspection failure. Recognizes limitations of mental state attribution abilities.

27.8.4 Constraint Parameters:

Abilities include constraint conditions, allowing for individual differences and context sensitivity in theory of mind capabilities.

27.8.5 Group vs. Individual:

Distinction between general mental models and group-specific models captures both universal and stereotype-based social cognition.

27.8.6 Belief-Based Models:

Mental models are beliefs in causal systems rather than direct mental access. Maintains epistemic limitations while enabling predictive reasoning.

27.9 Theoretical Significance:

Chapter 27 provides a concise but comprehensive formal treatment of theory of mind that integrates seamlessly with the broader cognitive architecture established in previous chapters.

The chapter's theoretical strength lies in its parsimony - rather than introducing separate machinery for theory of mind, it shows how mental state reasoning emerges naturally from the general causal reasoning framework. This integration explains why theory of mind abilities correlate with general reasoning abilities and why mental state attribution follows similar patterns to other causal attribution.

The treatment of introspection as a special case of mental state reasoning (where agent = target) provides insight into why self-knowledge is sometimes difficult and can fail. This challenges privileged access accounts of self-knowledge while maintaining the phenomenological reality that we have some access to our own mental processes.

The formal distinction between individual and group mental models provides foundation for understanding both empathic perspective-taking and social stereotyping within a unified framework. This integration is crucial for understanding how social cognition operates across different social contexts.

The constraint-based ability framework captures the reality that theory of mind abilities vary across individuals and contexts while maintaining the formal precision needed for computational implementation. This enables modeling of developmental differences, individual variation, and situational factors in mental state reasoning.

The minimal use of defeasible reasoning (only 1 axiom with etc) reflects that theory of mind concepts have relatively clear structural definitions, though their application depends on defeasible inferences about others' mental states embedded in the causal reasoning process.

This chapter establishes the formal foundation for social cognition, providing the infrastructure needed for modeling communication, cooperation, deception, empathy, and other phenomena requiring mental state attribution. It demonstrates how sophisticated social cognitive abilities can emerge from general-purpose causal reasoning mechanisms applied to the special case of mental causation.

27.10 Pattern Analysis:

- **6 Definitions** - Core framework for theory of mind concepts
- **1 Defeasible Rule** - Agent abilities with constraint conditions

- **2 Existence Claims** - Mental models of individuals and groups

The chapter represents an elegant extension of the causal reasoning framework to social cognition, maintaining theoretical integration while capturing the essential features of theory of mind and introspection.

Chapter 28

Chapter 28: Goals

28.1 Overview

- **82 axioms total** covering goals, subgoals, plans, trying/succeeding/failing, functionality, value/cost/importance, and multi-agent goal interactions
- **7 main sections**: Goals/Subgoals/Plans, Content of Goals, Goals and Multiple Agents, Trying/Succeeding/Failing, Functionality, Good and Bad, Value/Cost/Importance
- **Pure psychology domain** - all axioms deal with intentional agency and goal-directed behavior

28.2 Key Features Identified:

28.2.1 1. Foundational Goal Theory:

- **Goal Definition (28.1)**: Type constraint - goals are eventualities that agents have
- **Causal Knowledge (28.2-28.3)**: Agents know facts about what causes or enables what in the world
- **Goal Adoption (28.4-28.8)**: If agent has goal e2 and believes e1 causes/enables e2, agent adopts e1 as subgoal
- **Defeasible Planning**: Only axiom 28.8 uses defeasible reasoning with (etc) for goal adoption

28.2.2 2. Subgoal Structure and Planning:

- **Causation vs. Enablement**: Separate treatment for causal and enabling conditions in planning
- **Subgoal Definition (28.13)**: Subgoals are members of causal complexes for supergoals that agents believe in
- **Transitivity (28.14)**: Subgoal relation is transitive, allowing hierarchical goal structures

- **Goal Reversal**: Predicate goal reverses causality - wanting light on causes wanting to flip switch

28.2.3 3. Temporal Goal Categories:

- **Knowledge Goals (28.16)**: Goals of knowing something - prerequisite for most actions
- **Preservation Goals (28.17-28.18)**: Maintaining states over time intervals, with violation conditions
- **Persistent Goals (28.19)**: Goals that remain even after achievement (wealth, shell collecting)
- **Future Goals (28.20)**: Envisioned goals for future times (retirement planning)
- **Achievement States (28.21-28.23)**: Achieved, unachieved, and never-achieved goal classifications

28.2.4 4. Goal Interactions and Conflicts:

- **Conflicting Goals (28.24)**: Goals that cannot both be achieved simultaneously
- **Auxiliary Goals (28.25)**: Goals abandoned when conflicts arise with primary goals
- **Goal Hierarchy**: Sophisticated priority system for resolving goal conflicts

28.2.5 5. Multi-Agent Goal Theory:

- **Shared Goals (28.26-28.27)**: Collective goals with mutual belief among group members
- **Competitive Goals (28.28)**: Multiple agents wanting same property for themselves (races, competitions)
- **Adversarial Goals (28.29)**: Direct opposition where one agent wants negation of other's goal
- **Goal Attribution**: Framework for agents to understand and predict others' goals

28.2.6 6. Action Theory - Trying, Succeeding, Failing:

- **Trying Definition (28.30)**: Executing subgoals that are actions, caused by having those subgoals
- **Causal Involvement (28.31)**: Trying implies believing actions are causally involved in goal achievement
- **Success (28.32)**: Trying that causes the goal to actually occur
- **Failure (28.33)**: Trying without goal occurrence - allows for "lucking out" vs. genuine success

28.2.7 7. Functionality Theory:

- **Absolute Functionality (28.34)**: Composite entities associated with hypothetical agent goals
- **Relative Functionality (28.35)**: Component behaviors serving whole-system functionality
- **Intactness (28.36)**: Components able to fulfill functionality without impediments

- **Goal Talk Extension**: Framework applies to artifacts, organizations, and natural systems

28.2.8 8. Value, Cost, and Importance Framework:

- **Good/Bad Definitions (28.37-28.38)**: Events good/bad based on causal contribution to goals
- **Partial Orderings (28.39-28.44)**: Transitive comparison relations for value, cost, importance
- **Supergoal Hierarchy (28.45-28.52)**: Upper bound and least upper bound supergoals determine value rankings
- **Goal Relevance (28.53-28.58)**: Positive, negative, and general goal relevance for events and consequences

28.2.9 9. Scale Theory Integration:

- **Value/Cost/Importance Scales (28.74-28.82)**: Formal scale definitions using partial orderings
- **Hi Region Classification**: Valuable, costly, important as Hi region membership
- **Scale Positions**: Value, cost, importance as positions within respective scales
- **Property-Based Transfer**: Entity value/cost/importance derived from goal-relevant properties

28.3 Technical Sophistication:

28.3.1 Extensive Reification:

- 21 different primed predicates for goal processes: `goal'`, `cause'`, `enable'`, `believe'`, `subgoal'`, `try'`, `succeed'`, `moreValuable'`, etc.
- Enables sophisticated temporal and causal reasoning about intentional processes

28.3.2 Scale Theory Integration:

- Deep integration with Chapter 12 scale theory through `scaleDefinedBy`, `Hi`, `inScale` predicates
- Formal partial orderings for value, cost, and importance comparisons
- Systematic transfer principles from goals to entities via properties

28.3.3 Minimal Defeasible Reasoning:

- Only 2 axioms use `(etc)` conditions (28.8, 28.15) - focus on precise goal relationships
- Most goal reasoning follows deterministic patterns once beliefs are established

28.3.4 Multi-Agent Sophistication:

- Complex treatment of shared, competitive, and adversarial goals

- Mutual belief integration for collective agency
- Goal attribution and other-agent reasoning foundations

28.3.5 Temporal Complexity:

- Sophisticated temporal logic with `atTime`, `atTime'`, `before`, `during`, `during'`
- Time-anchored goal states, achievement conditions, and preservation requirements
- Future-directed planning and expectation integration

28.4 Complexity Distribution:

- **Simple: 20 axioms** (type constraints, basic goal reasoning, transitivity properties)
- **Moderate: 48 axioms** (standard goal definitions, multi-agent interactions, scale operations)
- **Complex: 14 axioms** (sophisticated temporal reasoning, functionality theory, upper bound supergoals)

28.5 Conceptual Importance:

28.5.1 Cognitive Architecture:

Provides comprehensive formal foundation for intentional agency. Shows how goals drive planning, action selection, and value judgments through systematic causal reasoning and belief integration.

28.5.2 AI Planning Systems:

Offers implementable framework for goal-directed reasoning with hierarchical planning, conflict resolution, and multi-agent coordination. Integrates trying/succeeding/failing for robust plan execution.

28.5.3 Philosophy of Mind:

Addresses fundamental questions about intentionality, agency, and value. Formalizes relationships between goals, actions, and evaluative judgments while maintaining computational tractability.

28.5.4 Multi-Agent Systems:

Sophisticated framework for shared goals, competition, and adversarial interactions. Provides foundation for understanding cooperation, conflict, and coordination in multi-agent environments.

28.5.5 Value Theory:

Formal account of how goals ground evaluative judgments about value, cost, and importance. Links subjective preferences to objective causal structures through goal-relevant properties.

28.6 Cross-Chapter Connections:

- **Chapter 15 (Causality)**: Uses `cause`, `enable`, `causalComplex` as foundation for goal reasoning
- **Chapter 21 (Belief Management)**: Uses `believe` for goal adoption and plan formation
- **Chapter 24 (Envisioning)**: Uses `causallyInvolved` for action planning and trying
- **Chapter 12 (Scales)**: Deep integration for value, cost, and importance measurements
- **Chapter 19 (Persons)**: Uses `agent` and `person` predicates for intentional subjects
- **Temporal Framework**: Integrates with temporal logic from earlier chapters

28.7 Applications Mentioned:

- **Hierarchical Planning**: Subgoal decomposition for complex task achievement
- **Multi-Agent Coordination**: Shared goals in organizations (General Motors selling cars)
- **Competitive Scenarios**: Racing, elections, market competition with goal conflicts
- **Artifact Design**: Functionality analysis of cars, steering wheels, complex systems
- **Natural Systems**: Functional analysis of trees, volcanos through hypothetical agency
- **Value Judgments**: Economic decisions, importance ranking, cost-benefit analysis

28.8 Notable Design Decisions:

28.8.1 Goal-Belief Integration:

Goals adopted through belief in causal/enabling relations rather than direct causal facts. Enables planning with false beliefs and accommodates individual differences in causal understanding.

28.8.2 Defeasible Minimalism:

Limited use of defeasible reasoning reflects focus on structural relationships rather than probabilistic inference. Goal adoption is defeasible, but goal definitions are precise.

28.8.3 Functionality Generalization:

Extension of goal framework to artifacts and natural systems through hypothetical agents. Enables unified treatment of intentional design and functional analysis.

28.8.4 Scale-Based Evaluation:

Value, cost, and importance grounded in formal scale theory rather than utility functions. Maintains qualitative reasoning while enabling comparative judgments.

28.8.5 Multi-Agent Integration:

Sophisticated treatment of collective agency, competition, and conflict. Mutual belief integration for shared goals while maintaining individual goal autonomy.

28.8.6 Temporal Anchoring:

Goals, achievements, and preservation anchored to specific times enabling precise temporal reasoning and plan monitoring.

28.8.7 Action-Goal Connection:

Trying defined through subgoal execution with causal requirements. Distinguishes genuine success from accidental achievement through causal role of intentions.

28.9 Theoretical Significance:

Chapter 28 represents the most comprehensive formal treatment of goal theory in cognitive science, integrating planning, action, evaluation, and multi-agent coordination into a unified framework.

The chapter's strength lies in its systematic integration across multiple levels: from individual goal adoption through belief and causation, to hierarchical planning with subgoal structures, to multi-agent scenarios with shared and conflicting goals, to evaluative frameworks grounding value judgments in goal-relevant properties.

The minimal use of defeasible reasoning (only 2/82 axioms) reflects the focus on structural relationships in goal systems rather than probabilistic planning. Once beliefs about causation and enabling are established, goal adoption and planning follow deterministic patterns, though the beliefs themselves may be defeasible.

The functionality theory provides an elegant extension of intentional concepts to artifacts and natural systems by associating them with hypothetical agents. This enables functional analysis while maintaining the goal-theoretic foundation and avoiding problematic teleological commitments.

The scale theory integration (28.74-28.82) grounds evaluative concepts in formal comparative structures rather than utility functions. This maintains the qualitative character of commonsense evaluation while enabling systematic reasoning about value, cost, and importance relationships.

The multi-agent framework addresses fundamental issues in collective agency, competition, and cooperation. The treatment of shared goals through mutual belief provides foundation for understanding organizational behavior, while competitive and adversarial goal frameworks address conflict and coordination challenges.

The temporal sophistication enables precise reasoning about goal achievement, preservation, and planning over time. The distinction between achieved goals (causally successful), unachieved goals

(attempted but failed), and never-achieved goals (never attempted) provides fine-grained analysis of success and failure patterns.

The trying/succeeding/failing framework addresses classic issues in action theory about the relationship between intentions and outcomes. The requirement that trying involve executing subgoals that are caused by having those subgoals ensures that genuine agency involves appropriate causal connections between goals and actions.

This chapter establishes the formal foundation for intentional agency that underlies rational choice, planning systems, multi-agent coordination, and evaluative judgment. It demonstrates how sophisticated intentional phenomena can emerge from systematic application of causal reasoning, belief integration, and hierarchical goal structures.

28.10 Pattern Analysis:

- **50 Definitions** - Comprehensive formalization of goal theory concepts
- **23 Goal Reasoning** - Systematic inference patterns for goal adoption and evaluation
- **4 Type Constraints** - Structural requirements for goal predicates
- **3 Argument Structure** - Transitivity and structural properties
- **2 Defeasible Rules** - Limited non-monotonic reasoning for goal adoption
- **2 Axiom Schema** - General patterns for causal knowledge

The chapter represents the culmination of the commonsense psychology framework, showing how abstract theories of causation, belief, time, and scales support sophisticated intentional agency suitable for both theoretical understanding and computational implementation.

Chapter 29

Chapter 29: Goal Themes

29.1 Overview

- **34 axioms total** covering thriving as ultimate goal, pleasure and pain theory, short-term vs long-term goal conflicts, and goal themes for group identity
- **4 main sections**: Thriving, Pleasure and Pain, Short-term vs Long-term Goals, Goal Themes
- **Mixed domains** - 31 psychology axioms (individual and group psychology) + 3 example axioms (formal illustration)

29.2 Key Features Identified:

29.2.1 1. Foundational Thriving Theory:

- **Universal Goal (29.1-29.2)**: Every agent has thriving as their ultimate goal - formal convenience for unified planning
- **Cultural Relativity**: Thriving content varies across cultures and individuals but provides universal structural framework
- **Planning Foundation**: Reduces mysterious goal emergence to planning problems about what causes thriving
- **Examples**: "Live long and prosper" vs. self-sacrifice for group, individual survival vs. collective thriving

29.2.2 2. Sophisticated Pleasure Theory:

- **Experiential Definition (29.3-29.7)**: Pleasure as private, self-evident experience with focused attention
- **Causal Structure (29.8)**: Pleasurable events cause pleasure experiences through explicit causal relations

- **Knowledge Properties (29.4-29.6, 29.9)**: Pleasure is known when experienced, believed pleasure is real, causes may be unknown
- **Goal Formation (29.10)**: Principal property - believing something pleasurable causes goal adoption
- **Thriving Connection (29.11)**: Pleasure causes belief in thriving, linking hedonic experience to ultimate goal
- **Value Correlation (29.12)**: More valuable goals produce more pleasurable achievements

29.2.3 3. Comprehensive Pain Theory:

- **Parallel Structure (29.13-29.17)**: Pain mirrors pleasure theory - private, self-evident, causally structured
- **Body Damage Indication (29.19)**: Pain typically signals damage to body parts through changeFrom intact states
- **Avoidance Goals (29.20)**: Painful things cause goals to stop the pain - opposite of pleasure's approach goals
- **Thriving Impact (29.21)**: Pain causes belief in not thriving, contrasting with pleasure's positive effect
- **Cost Correlation (29.22)**: Pain intensity correlates with costs agents willing to sacrifice for relief

29.2.4 4. Temporal Choice Framework:

- **Short-term vs Long-term (29.23)**: Formal framework for temporal goal conflicts with strict constraints
- **Conflict Requirements**: Goals must conflict (can't have both), long-term must be more valuable, temporal ordering
- **Decision Mechanisms (29.24-29.25)**: Picking short/long-term options involves abandoning the alternative goal
- **Gratification Delay (29.26-29.27)**: Time interval between satisfactions has inherent costs
- **Real-world Applications**: Bread before dinner, job vs. medical school, exercise pain tolerance

29.2.5 5. Group Psychology - Goal Themes:

- **Mutual Belief Integration (29.28-29.30)**: Knowledge tagged by group membership through mutual belief structures
- **Goal Theme Definition (29.31)**: Sets of goals associated with agent groups enabling behavioral prediction
- **Causal Grounding (29.32)**: Goal themes based on mutual belief about what causes thriving for group members

- **Identity Formation (29.33)**: Self-identity expressed through goal theme membership ("As an American, I vote")
- **Hedonist Example (29.34)**: Formal definition of hedonists as those whose goal themes involve pleasurable experiences

29.3 Technical Sophistication:

29.3.1 Extensive Defeasible Reasoning:

- **10 axioms use (etc)** conditions - highest proportion in psychology chapters so far
- Reflects inherently uncertain nature of pleasure/pain responses, goal adoption, and group identity
- Captures individual variation in hedonic responses and cultural differences in goal themes

29.3.2 Reification Complexity:

- **13 different primed predicates**: `thrive'`, `pleasure'`, `pain'`, `pleasurable'`, `painful'`, `believe'`, `changeFrom'`, `atTime'`, `pickST'`, `pickLT'`, `imply'`, `causallyInvolved'`
- Enables sophisticated reasoning about temporal processes, hedonic states, and group belief structures

29.3.3 Multi-level Integration:

- **Individual Psychology**: Pleasure/pain experiences, goal formation, temporal choice
- **Group Psychology**: Mutual belief, goal themes, cultural identity, stereotyping
- **Temporal Reasoning**: Before relations, gratification delays, short/long-term conflicts

29.3.4 Cultural Framework:

- **Universal Structure**: All agents have thriving goal with formal planning processes
- **Cultural Content**: Thriving definitions vary (survival, pleasure, group welfare, spiritual goals)
- **Group Differentiation**: Academic, religious, national, professional goal themes

29.4 Complexity Distribution:

- **Simple: 5 axioms** (type constraints, basic belief properties, example schemas)
- **Moderate: 16 axioms** (standard definitions, defeasible rules, causal structures)
- **Complex: 13 axioms** (temporal choice framework, goal theme integration, multi-level causation)

29.5 Conceptual Importance:

29.5.1 Ultimate Goal Theory:

Provides formal foundation for unified agency by positing thriving as universal goal while maintaining cultural relativity in content. Solves the mystery of goal origin by reducing it to planning problems about causal beliefs.

29.5.2 Hedonic Psychology:

Sophisticated formal treatment of pleasure and pain that explains their roles in goal formation and thriving assessment. Links evolutionary mechanisms (rough and ready pleasure/pain systems) to rational goal structures.

29.5.3 Temporal Psychology:

Formal framework for understanding self-control, delayed gratification, and temporal preference conflicts. Captures both structural requirements (conflict, value ordering) and process mechanisms (goal abandonment, delay costs).

29.5.4 Group Identity Theory:

Groundbreaking integration of individual goal psychology with group identity through mutual belief and goal themes. Provides formal foundation for cultural psychology, stereotyping, and social prediction.

29.5.5 Cross-Cultural Framework:

Balances universal cognitive architecture (thriving goal, planning mechanisms) with cultural relativity (thriving content, goal themes). Enables formal comparison across cultures while respecting differences.

29.6 Cross-Chapter Connections:

- **Chapter 28 (Goals)**: Uses `goal`, `moreValuable`, `moreCostly`, `conflictingGoals` as foundation for hedonic and temporal choice
- **Chapter 21 (Belief Management)**: Uses `believe'`, `know`, `inFocus` for conscious access to hedonic states
- **Chapter 21 (Mutual Belief)**: Uses `mb` predicate for goal theme and group identity structures
- **Chapter 15 (Causality)**: Uses `cause` extensively for pleasure/pain causation and goal formation
- **Temporal Framework**: Integrates `before`, `atTime'` for temporal choice and gratification delay

- **Chapter 24 (Envisioning)**: Uses `causallyInvolved'` for goal theme justification

29.7 Applications Mentioned:

- **Cultural Examples**: Spock's "Live long and prosper", suicide bombers, soldiers, accident rescuers
- **Hedonic Examples**: Decadent desserts, "no pain no gain" exercise, phantom limb pain
- **Temporal Examples**: Bread before dinner, job vs. medical school, ant and grasshopper fable
- **Group Examples**: Academic beliefs, American democracy, computer science knowledge, religious faith
- **Identity Examples**: "As an American, I vote", professional role expectations, lifestyle choices

29.8 Notable Design Decisions:

29.8.1 Thriving Universality:

Formal convenience assumption that all agents have thriving as ultimate goal. Provides unified planning framework while allowing cultural variation in thriving content.

29.8.2 Hedonic Causality:

Explicit causal connections between events, pleasure/pain experiences, and goal formation. Captures both approach (pleasure $\rightarrow$ goals) and avoidance (pain $\rightarrow$ stop goals) motivations.

29.8.3 Temporal Structure:

Strict requirements for short/long-term conflicts (must conflict, value ordering, temporal sequence). Ensures framework applies only to genuine temporal choice dilemmas.

29.8.4 Group-Individual Integration:

Goal themes connect group identity to individual goal adoption through mutual belief structures. Enables prediction from group membership while maintaining individual agency.

29.8.5 Cultural Relativity:

Universal cognitive architecture (thriving goal) with culturally specific content. Avoids both cognitive relativism and cultural imperialism through structural/content distinction.

29.8.6 Defeasible Dominance:

Heavy use of defeasible reasoning (10/34 axioms) reflects uncertain, context-dependent nature of hedonic responses and cultural identity effects.

29.9 Theoretical Significance:

Chapter 29 represents a major integration of individual psychology (hedonic experience, temporal choice) with group psychology (cultural identity, mutual belief) under a unified goal-theoretic framework.

The thriving theory provides elegant solution to goal emergence problem by positing universal structural goal while allowing cultural variation in content. This enables formal cross-cultural psychology without imposing specific values or losing predictive power.

The pleasure/pain theory offers sophisticated alternative to simple hedonism by grounding hedonic responses in causal structures and connecting them to ultimate thriving goal. The evolutionary perspective (rough and ready mechanisms) combined with rational choice (goal formation, value correlation) captures both automatic and deliberative aspects of hedonic psychology.

The temporal choice framework provides first formal treatment of self-control and delayed gratification that captures both structural requirements (goal conflict, value ordering) and process mechanisms (goal abandonment, delay costs). The gratification delay concept with inherent costs explains why temporal choice is difficult even when long-term goals are more valuable.

The goal theme theory represents breakthrough in formal social psychology by connecting group identity to individual behavior through mutual belief and causal reasoning. The framework explains how group membership enables behavioral prediction while maintaining individual cognitive autonomy.

The extensive defeasible reasoning (10/34 axioms) reflects deep appreciation for context-dependency in psychology. Hedonic responses vary across individuals and situations, cultural identity effects have exceptions, and temporal choices depend on complex situational factors.

The integration across levels (individual experience, goal formation, group identity) demonstrates how sophisticated social psychological phenomena emerge from systematic application of belief, causation, and goal structures established in earlier chapters.

This chapter establishes the foundation for understanding human motivation as emerging from universal cognitive architecture (thriving-oriented planning) operating with culturally specific content (goal themes, hedonic values) through both individual (pleasure/pain, temporal choice) and social (group identity, mutual belief) mechanisms.

29.10 Pattern Analysis:

- **12 Definitions** - Formal characterization of thriving, hedonic states, temporal choice, goal themes
- **10 Defeasible Rules** - Non-monotonic patterns in pleasure/pain responses, goal formation, cultural identity

- **4 Belief Logic** - Conscious access to hedonic states and self-evidence properties
- **4 Axiom Schema** - Templates for cultural belief systems and group identity rules
- **3 Type Constraints** - Basic structural requirements for thriving, pleasure, pain predicates
- **2 Argument Structure** - Privacy conditions for hedonic experiences
- **1 Each**: Existence Claim (universal thriving goal), Goal Reasoning (theme justification)

The chapter represents the culmination of individual psychology integration with social psychology, showing how personal experience (hedonic states, temporal choice) connects to cultural identity (goal themes, group membership) through sophisticated belief and causation structures that maintain both universal cognitive architecture and cultural specificity.

Chapter 30

Chapter 30: Threats and Threat Detection

30.1 Overview

- **23 axioms total** covering threat situations, threat detection, risk bias, concern, and seriousness evaluation
- **3 main sections**: Threat Situations, Threat Detection and Management, Seriousness
- **Pure psychology domain** - all axioms deal with cognitive threat assessment and risk management

30.2 Key Features Identified:

30.2.1 1. Foundational Threat Theory:

- **Threat Situation Definition (30.1)**: Complex framework using envisioned causal systems (ecs) with branch reasoning
- **Branching Structure**: Threats arise when one branch (e1) leads to goal negation while alternative (e2) does not
- **Agent-Relative Assessment**: Threats always defined relative to agent's goals and causal understanding
- **ECS Integration**: Deep connection to Chapter 24 envisioning framework through causal systems

30.2.2 2. Modal Threat Categories:

- **Real vs. Possible vs. Impossible (30.6-30.8)**: Systematic modal distinctions based on constraints and world understanding

- **Real Threats**: When envisioned causal system is current world understanding and anchor really exists
- **Possible Threats**: Real threat situations where threat outcome is possible given constraints
- **Impossible Threats**: Real threat situations where threat outcome violates constraints
- **Temporal Realization (30.9-30.10)**: Realized vs. unrealized threats based on temporal occurrence

30.2.3 3. Cognitive Threat Processing:

- **Threat Detection (30.11-30.13)**: Process of identifying threat situations through envisioning sequences
- **Detection Abilities**: People have inherent threat detection capabilities but can fail
- **Concern Framework (30.14)**: Being concerned involves believing in possible threats to relevant entities
- **Threat Management (30.17-30.18)**: Adding and removing threats from concern sets through belief change

30.2.4 4. Risk Bias Theory:

- **Bias Against Risk (30.15)**: When concerns about threats likely cause avoidance of risky actions
- **Bias Toward Risk (30.16)**: When concerns about threats do not likely cause avoidance of risky actions
- **Individual Differences**: Systematic personality-based differences in risk assessment and response
- **Causal Mechanisms**: Risk bias grounded in likelihood of concern-to-action causal chains

30.2.5 5. Seriousness Framework:

- **Composite Scale Theory (30.19)**: Seriousness as composite of likelihood and importance scales
- **Scale Integration**: Uses Chapter 12 scale theory with Chapter 20 likelihood and Chapter 28 importance
- **Comparative Seriousness (30.20)**: More serious relationships using scale-defined orderings
- **Absolute Seriousness (30.21)**: Hi region membership in seriousness scale
- **Threat Assessment (30.23)**: Process of trying to determine threat seriousness levels

30.3 Technical Sophistication:

30.3.1 Extensive Reification:

- **11 primed predicates**: `not'`, `realThreatSituation'`, `possibleThreat'`, `concern'`, `cause'`, `changeTo'`, `changeFrom'`, `know'`, `moreSerious'`, `serious'`, `seriousness'`
- Enables sophisticated temporal and causal reasoning about threat processes
- Supports meta-cognitive reasoning about threat assessment and management

30.3.2 Temporal Logic Integration:

- **Temporal predicates**: `atTime`, `before` for precise temporal relationships
- **Threat Realization**: Temporal constraints on when threats were anticipated vs. when they occurred
- **Dynamic Processes**: Threat detection, assessment, and management as temporally extended processes

30.3.3 Envisioned Causal Systems Framework:

- **Deep ECS Integration**: Uses `ecs`, `ecsSeq`, `causallyInvolved`, `branch` from Chapter 24
- **Branch Reasoning**: Systematic treatment of alternative futures and their implications
- **World Understanding**: Grounding in agent's current world understanding (`cwu`)

30.3.4 No Defeasible Reasoning:

- **Unique Precision**: No axioms use `(etc)` conditions - focus on exact threat definitions
- **Mathematical Rigor**: Precise logical relationships without defeasible exceptions
- **Foundational Framework**: Establishes firm definitions for more flexible applications

30.3.5 Scale Theory Integration:

- **Composite Scales**: Sophisticated integration of likelihood and importance into seriousness
- **Multi-dimensional Assessment**: Threats evaluated along multiple dimensions simultaneously
- **Formal Comparisons**: Scale-based ordering relationships for threat prioritization

30.4 Complexity Distribution:

- **Simple: 7 axioms** (basic definitions, agent abilities, threat categorization)
- **Moderate: 12 axioms** (modal threat distinctions, detection processes, seriousness framework)
- **Complex: 4 axioms** (core threat situation definition, risk bias theory, temporal realization)

30.5 Conceptual Importance:

30.5.1 Cognitive Architecture:

Provides comprehensive formal foundation for threat perception, assessment, and response. Shows how threat cognition emerges from integration of causal reasoning, temporal logic, modal concepts, and evaluative scales.

30.5.2 AI Safety and Security:

Formal framework for threat modeling in autonomous systems. Enables systematic threat detection, assessment, and response planning with clear modal distinctions and risk evaluation.

30.5.3 Psychology of Risk:

Addresses fundamental questions about individual differences in risk perception and response. Formalizes relationships between threat assessment, concern, and behavioral bias patterns.

30.5.4 Decision Theory:

Links threat assessment to goal-based decision making through formal seriousness evaluation and risk bias mechanisms. Provides foundation for rational threat response under uncertainty.

30.5.5 Temporal Reasoning:

Sophisticated treatment of threats as temporally extended phenomena with anticipation, realization, and assessment phases requiring precise temporal logic.

30.6 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Uses eventuality framework and `Rexist` for threat realization
- **Chapter 12 (Scales)**: Deep integration with scale theory for seriousness evaluation
- **Chapter 15 (Causality)**: Uses causal relations for threat situation definitions
- **Chapter 19 (Persons)**: Uses `person` and `agent` predicates for threat subjects
- **Chapter 20 (Modality)**: Uses `possible` and `likely` for modal threat distinctions
- **Chapter 21 (Belief Management)**: Uses `believe` for concern and threat assessment
- **Chapter 24 (Envisioning)**: Fundamental dependency on ECS framework and branch reasoning
- **Chapter 26 (Managing Expectations)**: Uses `cwu` (current world understanding) for threat grounding
- **Chapter 28 (Goals)**: Uses `goal` predicate for threat target identification

30.7 Applications Mentioned:

- **Personal Safety**: Threat detection and avoidance in daily life situations
- **Risk Assessment**: Individual differences in risk perception and decision making
- **Security Analysis**: Systematic threat modeling and vulnerability assessment
- **Planning Systems**: Threat-aware planning with risk bias considerations
- **Behavioral Economics**: Risk bias effects on economic decision making

30.8 Notable Design Decisions:

30.8.1 ECS Foundation:

Threats grounded in sophisticated envisioned causal systems rather than simple probability assessments. Enables rich causal reasoning about threat scenarios and their implications.

30.8.2 Agent-Relative Definition:

All threat concepts defined relative to specific agents and their goals. Captures subjective nature of threat perception while maintaining formal precision.

30.8.3 Modal Sophistication:

Systematic treatment of real, possible, and impossible threats based on constraints and world understanding. Enables precise reasoning about threat feasibility.

30.8.4 Temporal Precision:

Exact temporal relationships between threat anticipation and realization. Supports learning from threat assessment accuracy and timing.

30.8.5 Composite Evaluation:

Seriousness as composite of likelihood and importance rather than simple probability. Captures multidimensional nature of threat assessment.

30.8.6 Risk Bias Integration:

Individual differences in risk response formalized through likelihood of concern-to-action causal chains. Enables personality-based threat response modeling.

30.8.7 No Defeasible Reasoning:

Precise definitions without defeasible exceptions establish firm foundation for threat reasoning. Reflects mathematical nature of threat logic.

30.9 Theoretical Significance:

Chapter 30 provides the most comprehensive formal treatment of threat cognition in cognitive science, integrating causal reasoning, temporal logic, modal concepts, and evaluative scales into a unified framework for threat assessment and response.

The chapter's strength lies in its systematic integration across multiple levels: from basic threat situation definitions through sophisticated envisioned causal systems, to modal distinctions between possible and impossible threats, to individual differences in risk bias and assessment.

The deep integration with the envisioned causal systems framework from Chapter 24 grounds threat reasoning in sophisticated causal understanding rather than simple statistical patterns. This enables rich scenario planning and counterfactual reasoning about threat mitigation strategies.

The temporal precision enables exact specification of when threats are anticipated, when they might occur, and when they are realized or fail to materialize. This mathematical rigor supports both theoretical analysis and practical implementation in AI systems requiring temporal threat reasoning.

The modal sophistication (real vs. possible vs. impossible threats) provides a principled framework for reasoning about threat feasibility under different constraint sets. This enables systematic threat prioritization and resource allocation for threat mitigation.

The composite seriousness evaluation (likelihood $\times$ importance) captures the intuitive multidimensional nature of threat assessment while maintaining formal precision. This provides foundation for rational threat prioritization and response planning.

The risk bias theory addresses fundamental individual differences in threat response through formal causal mechanisms. Rather than treating risk bias as personality traits, the framework grounds bias patterns in likelihood differences for concern-to-action causal chains.

The absence of defeasible reasoning (no `(etc)` conditions) reflects the foundational nature of threat definitions. While threat assessment applications may involve uncertainty and exceptions, the basic logical structure of threat situations follows precise patterns.

This chapter establishes threat cognition as emerging from systematic application of causal reasoning, temporal logic, modal concepts, and evaluative scales. It provides the foundation for understanding risk perception, security analysis, and threat-aware planning in both human and artificial agents.

30.10 Pattern Analysis:

- **23 Definitions** - Comprehensive formalization of threat theory concepts from basic situations to complex assessment processes
- **1 Existence Claim** - Basic agent capability for threat detection

The chapter represents a focused, mathematically rigorous treatment of threat cognition that

maintains deep integration with the broader causal reasoning, temporal logic, and evaluative frameworks while providing precise tools for modeling risk perception and threat response in uncertain environments.

30.11 Unique Contributions:

30.11.1 Branch-Based Threat Logic:

Novel use of branch reasoning from envisioned causal systems to define threats as goal-threatening alternatives in causal scenarios. This provides richer threat modeling than simple probability-based approaches.

30.11.2 Modal Threat Taxonomy:

Systematic classification of threats by modal status (real/possible/impossible) and temporal realization (realized/unrealized) creates comprehensive framework for threat categorization and assessment.

30.11.3 Composite Seriousness Theory:

Integration of likelihood and importance scales into composite seriousness evaluation provides formal foundation for multidimensional threat assessment that captures both probability and value considerations.

30.11.4 Risk Bias Formalization:

Individual differences in risk response formalized through likelihood of concern-to-action causal chains rather than personality traits, enabling systematic modeling of risk behavior patterns.

30.11.5 Temporal Threat Dynamics:

Precise temporal logic for threat anticipation, realization, and assessment enables formal analysis of threat prediction accuracy and timing effects on decision making.

30.11.6 ECS-Grounded Threat Reasoning:

Integration with envisioned causal systems framework provides sophisticated foundation for threat scenario analysis, counterfactual reasoning, and mitigation planning that goes beyond statistical threat models.

This chapter establishes threat cognition as a sophisticated integration of causal understanding, temporal reasoning, modal logic, and evaluative judgment that supports both human psychological understanding and computational implementation in AI systems requiring robust threat assessment capabilities.

Chapter 31

Chapter 31: Plans

- **53 axioms total** covering plans as mental entities, planning processes, strategies and tactics, executability, plan types, and helping
- **6 main sections**: Plans as Mental Entities, Planning Process, Strategies and Tactics, Executability and Complete Plans, Types of Plans, Helping

- **Pure psychology** - comprehensive treatment of planning as central cognitive process connecting beliefs to actions

31.1 Key Features Identified:

31.1.1 1. Plans as Mental Constructs (Axioms 31.1-31.9):

- **Axiom 31.1**: `connectedSubgoal*` - recursive transitive closure of subgoal relations ensuring plan connectivity
- **Axiom 31.2**: Plans as composite entities with components (subgoals) and relations (subgoal relations)
- **Axioms 31.3-31.4**: Convenience predicates for subgoal membership and plan component access
- **Axiom 31.5**: Subplans as recursive plan structures with subset component and relation requirements
- **Axiom 31.6**: Plan sequences capturing plan evolution through `change` relations between successive versions
- **Axiom 31.7**: Subgoal relations grounded in agent's causal beliefs about enablement/causation
- **Axioms 31.8-31.9**: Plans as envisioned causal systems, planning as variety of envisioning

31.1.2 2. Belief System vs. Planning Module Architecture (Axioms 31.10-31.22):

- **Axiom 31.10**: Desires as causal beliefs about what contributes to thriving (belief system)

- **Axioms 31.11-31.13**: `thePlan` as current plan for thriving, `thePlanseq` as temporal sequence of plans
- **Axiom 31.12**: Actions defeasibly caused by being subgoals in thePlan (using `etc` for non-monotonicity)
- **Axiom 31.14**: Intentions as action subgoals in thePlan that agent performs
- **Axioms 31.15-31.16**: `decideTo` as bridge from belief system to planning module via plan modification
- **Axiom 31.17**: `planTo` combining goal reasoning, plan sequence generation, and plan adoption
- **Axioms 31.18-31.22**: Will as mind-body interface with constraints, direct causation, and plan-based motivation

31.1.3 3. Strategies and Tactics Framework (Axioms 31.23-31.26):

- **Axioms 31.23-31.24**: Both strategies and tactics as parameterized plans (eventuality types with non-null parameter sets)
- **Axiom 31.25**: Hierarchical constraint - strategies not subplans of tactics
- **Axiom 31.26**: Defeasible rule that tactics serve strategic goals (subplans of strategies)
- **Distinction criteria**: Strategies longer-term/higher-level, tactics shorter-term/lower-level, tactical responses to immediate situations

31.1.4 4. Executability Theory (Axioms 31.27-31.30):

- **Axiom 31.27**: `executable1` extending basic executability to include physics and committed agents
- **Four executability sources**: Agent capability, natural occurrence, committed others, recursive subgoal executability
- **Axiom 31.28**: Commitment theory - defeasible guarantee that committed agents cause their assigned events
- **Axiom 31.29**: Complete plans defined by `executable1` top-level goals at current time
- **Axiom 31.30**: Partial plans as non-complete plans requiring further development

31.1.5 5. Plan Type Taxonomy (Axioms 31.31-31.49):

31.1.5.1 Basic Plan Types:

- **Axioms 31.31-31.32**: `planInstance` and `planInstancePlus` for type instantiation and elaboration
- **Axiom 31.33**: Normal plans as frequent/socially expected plan types

31.1.5.2 Multi-Agent Plan Relations:

- **Axiom 31.34**: Adversarial plans targeting negation of other agents' goals
- **Axiom 31.35**: Counterplans blocking adversarial plans through causal negation

- **Axiom 31.36-31.37**: Competitive goals/plans where mutual exclusivity creates side-effect opposition
- **Axiom 31.38**: Assistive plans adopting others' goals through causal goal transmission
- **Axiom 31.39**: Shared plans requiring group goal ownership, mutual belief in structure, and member commitment

31.1.5.3 Cognitive Accessibility:

- **Axiom 31.40**: Envisioned plans as causal systems modeled by other agents
- **Axiom 31.41**: Unknown plans not accessible to other agents
- **Axiom 31.42**: Nonconscious plans not accessible to planning agent themselves

31.1.5.4 Execution Patterns:

- **Axioms 31.43-31.44**: `do` for individual actions and entire plan execution
- **Axiom 31.45**: Reused plans through multiple distinct executed instances
- **Axioms 31.46-31.47**: Occasional and repeated plans as reused plan subtypes
- **Axiom 31.48**: Periodic plans with equal intervals between execution instances
- **Axiom 31.49**: Continuous plans executed over specified time intervals

31.1.6 6. Helping Theory (Axioms 31.50-31.53):

- **Axiom 31.50**: `help0` - basic causal involvement (unintentional helping)
- **Axiom 31.51**: `help1` - agent action in causal complex for another's goal (accidental helping)
- **Axiom 31.52**: `help2` - intentional action motivated by other's goal (deliberate helping)
- **Axiom 31.53**: `help3` - shared plan participation for other's goal (collaborative helping)

31.2 Technical Sophistication:

- **Recursive Definitions**: `connectedSubgoal*` and `executable1` with proper termination conditions
- **Composite Entity Theory**: Plans as structured collections of subgoals and subgoal relations
- **Extensive Reification**: 19 reified predicates for mental processes and plan relationships
- **Defeasible Reasoning**: 6 axioms using `(etc)` for non-monotonic planning and action
- **Multi-Agent Coordination**: Sophisticated treatment of competitive, adversarial, assistive, and shared planning
- **Cognitive Architecture**: Clear separation between belief system (desires, causal beliefs) and planning module (intentions, decisions, actions)

31.3 Complexity Distribution:

- **Simple**: 23 axioms (basic constraints, definitions, type relationships)

- **Moderate**: 21 axioms (standard planning definitions, multi-agent relations)
- **Complex**: 9 axioms (recursive definitions, plan sequences, shared plans, counterplans)

31.4 Conceptual Importance:

This chapter provides the central framework for: - **Action Generation**: Bridge from beliefs to actions through planning module - **Mental Architecture**: Dual-system model separating knowledge from goal-directed behavior
- **Multi-Agent Systems**: Comprehensive taxonomy of inter-agent plan relationships - **Cognitive Psychology**: Formal treatment of intentions, decisions, and goal-directed behavior - **AI Planning**: Executability theory, plan refinement, and social planning concepts - **Philosophy of Action**: Will as mind-body interface, conscious vs. nonconscious planning

31.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Uses eventuality framework and `Rexist` for plan component reality
- **Chapter 15 (Causality)**: Grounded in causal theory through `causallyInvolved` and `cause` relations
- **Chapter 21 (Belief)**: Belief system provides causal knowledge used by planning module
- **Chapter 24 (Envisioning)**: Plans as envisioned causal systems extending ECS theory
- **Chapter 28 (Goals)**: Goal theory provides foundation for plan target specification
- **Chapter 30 (Thriving)**: Ultimate goal of all planning, short-term vs. long-term considerations

31.6 Applications Mentioned:

- **Strategic Planning**: Customer loyalty through coupons, military strategy vs. tactics
- **Action Control**: Trapeze artist coordination, sofa moving cooperation
- **Multi-Agent Scenarios**: McCain-Obama election dynamics, competitive sports
- **Helping Behaviors**: Elderly street crossing, Christmas present scenarios, key confiscation
- **Plan Types**: Periodic exercise routines, continuous monitoring, occasional celebrations

31.7 Notable Design Decisions:

- **Composite Entity Framework**: Plans as structured objects rather than simple predicates
- **Dual Cognitive Architecture**: Clear separation between passive beliefs and active planning
- **Recursive Plan Structure**: Subplans allowing hierarchical plan organization
- **Social Plan Theory**: Extensive treatment of multi-agent planning relationships

- **Executability Integration**: Physics, agent capabilities, and social commitments in unified framework
- **Defeasible Planning**: Non-monotonic reasoning acknowledging plan modification and action variability
- **Will as Primitive**: Mind-body interface through willing rather than detailed neurophysiology

31.8 Theoretical Significance:

Chapter 31 represents the culmination of the commonsense psychology framework by connecting abstract mental representations (beliefs, desires, goals) to concrete physical actions. The sophisticated treatment of multi-agent planning provides formal foundations for social psychology and organizational behavior.

The dual architecture of belief system vs. planning module offers a cognitively plausible model that avoids both the frame problem (passive belief systems) and the symbol grounding problem (active planning with environmental interaction). The extensive reification enables precise reasoning about planning processes themselves.

The integration of individual cognitive processes with multi-agent social dynamics through shared plans, commitments, and helping relationships demonstrates how formal logic can capture both personal decision-making and interpersonal coordination. This provides foundations for both AI systems design and human social interaction modeling.

The chapter's 53 axioms establish planning as the central process transforming mental representations into physical actions, making it essential for any complete theory of agent behavior in both individual and social contexts.

Chapter 32

Chapter 32: Goal Management

- **20 axioms total** covering goal addition, removal, modification, priority theory, assessment, and execution control
- **3 main sections**: Adding/Removing/Modifying Goals, Priority, Assessing and Prioritizing Goals
- **Pure psychology** - sophisticated treatment of dynamic goal management as central cognitive process

32.1 Key Features Identified:

32.1.1 1. Goal Lifecycle Management (Axioms 32.1-32.9):

32.1.1.1 Goal Addition and Removal:

- **Axiom 32.1**: `addGoal` - adding goals through `changeTo` transitions in thePlan structure
- **Axiom 32.2**: `removeAchievedGoal` - goal removal caused by goal achievement in real world
- **Axiom 32.3**: `abandonGoal` - goal removal despite non-achievement (voluntary abandonment)
- **Axiom 32.4**: `removeViolatedGoal` - goal removal due to impossibility constraints (resignation)

32.1.1.2 Goal Modification:

- **Axiom 32.5**: `suspendGoal` - temporal postponement by shifting timing parameters from t1 to t2
- **Axiom 32.6**: `modifyGoal` - goal transformation requiring similarity between old and new goals
- **Axiom 32.7**: `specifyGoal` - modification to more specific goal via logical implication (g2 $\rightarrow$ g1)
- **Axiom 32.8**: `generalizeGoal` - modification to more general goal via logical implication (g1

$\rightarrow$ g2)

32.1.1.3 Goal Pursuit:

- **Axiom 32.9**: `pursueGoal` - three modes of goal pursuit (envisioning plans, deciding on plans, active trying)

32.1.2 2. Priority Theory Framework (Axioms 32.10-32.12):

32.1.2.1 Multi-Factor Priority Determination:

- **Axiom 32.10**: `priorityScale` - agent-specific scales for goal prioritization
- **Axiom 32.11**: **Complex defeasible priority rule** - importance, difficulty, and likelihood jointly determine priority
  - **Importance**: Higher importance $\rightarrow$ higher priority
  - **Difficulty**: Higher difficulty $\rightarrow$ lower priority (effort correlation)
  - **Likelihood**: Higher likelihood $\rightarrow$ higher priority
  - **Ceteris paribus conditions**: Other factors held constant for comparison
- **Axiom 32.12**: `priority` - goal priority as scale position

32.1.2.2 Priority Factors:

- **Importance**: From Chapter 28, constrained by subgoal relations
- **Effort/Difficulty**: Physical and mental energy expenditure, correlated with obstacles
- **Likelihood**: Probability of successful goal achievement

32.1.3 3. Goal Assessment and Conflict Resolution (Axioms 32.13-32.20):

32.1.3.1 Goal Justification and Assessment:

- **Axiom 32.13**: `goalJustification` - causal stories linking goals to higher-level goals through plans
- **Axiom 32.14**: `assessGoal` - envisioning causal chains from/to goals and their priorities

32.1.3.2 Conflict Management:

- **Axiom 32.15**: `conflict` - mutual temporal exclusivity of goals
- **Axiom 32.16**: `resolveConflictingGoals` - conflict resolution through goal modification

32.1.3.3 Preference and Prioritization:

- **Axiom 32.17**: `preferGoal` - preference based on goal properties causing preference for pursuit
- **Axiom 32.18**: `prioritizeGoal` - dynamic priority change through reified priority transitions

32.1.3.4 Priority-Based Execution Control:

- **Axiom 32.19**: **High priority → pursuit** - defeasible causation from high priority to goal pursuit
- **Axiom 32.20**: **Low priority → non-pursuit** - defeasible inhibition of goal pursuit from low priority

32.2 Technical Sophistication:

- **Dynamic Goal Structures**: Goals as mutable entities in constantly changing thePlan
- **Extensive Reification**: 11 reified predicates for goal states, priority changes, and mental processes
- **Scale Theory Integration**: Priority, importance, difficulty, and likelihood as scale positions
- **Defeasible Reasoning**: 4 axioms using `(etc)` for non-monotonic priority-based behavior
- **Temporal Manipulation**: Complex temporal parameter shifting in goal suspension
- **Multi-Factor Decision Making**: Sophisticated ceteris paribus reasoning for priority determination
- **Conflict Resolution**: Systematic approach to detecting and resolving goal conflicts

32.3 Complexity Distribution:

- **Simple**: 8 axioms (basic goal operations, constraint definitions, simple modifications)
- **Moderate**: 9 axioms (goal pursuit, priority definitions, assessment, conflict resolution)
- **Complex**: 3 axioms (goal suspension, priority ordering, low priority inhibition)

32.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Dynamic Planning**: Moving beyond static plans to continuously evolving goal structures - **Resource Allocation**: Priority-based attention and effort distribution among competing goals - **Cognitive Control**: Mechanisms for goal addition, modification, suspension, and abandonment - **Conflict Resolution**: Systematic approach to managing incompatible goals - **Rational Action**: Multi-factor decision making for goal prioritization and pursuit - **Temporal Planning**: Goal scheduling and temporal parameter manipulation

32.5 Cross-Chapter Connections:

- **Chapter 28 (Goals)**: Uses goal and importance predicates as foundation
- **Chapter 31 (Plans)**: Extends thePlan framework with dynamic goal management
- **Chapter 12 (Scales)**: Heavy integration with scale theory for priority, importance, difficulty, likelihood

- **Chapter 24 (Envisioning)**: Goal assessment through envisioning causal chains
- **Chapter 15 (Causality)**: Causal relations in goal achievement, removal, and priority effects
- **Chapter 21 (Belief)**: Agent model of world providing goal justifications
- **Chapter 37 (Planning Goals)**: Forward reference to preference theory

32.6 Applications Mentioned:

- **Goal Modification**: Vacation planning (Puerto Vallarta vs. Mazatlan), expensive vacation vs. new car
- **Goal Conflicts**: Reading novel vs. grocery shopping with no causal ordering
- **Priority Factors**: Short/easy jobs before long/hard ones, effort-based procrastination
- **Goal Abandonment**: Valuable vase protection after breakage (impossibility-based resignation)
- **Resource Allocation**: High priority goals get pursued, low priority goals get ignored/abandoned

32.7 Notable Design Decisions:

- **Similarity-Based Modification**: Goal changes require similarity relationships between old and new goals
- **Multi-Factor Priority**: Integration of importance, difficulty, and likelihood with ceteris paribus reasoning
- **Defeasible Execution**: Priority effects on goal pursuit are non-monotonic (exceptions possible)
- **Temporal Flexibility**: Goal suspension through temporal parameter manipulation rather than removal
- **Causal Justification**: Goals justified through causal stories linking to higher-level goals
- **Conflict Detection**: Systematic temporal exclusivity checking for goal compatibility
- **Dynamic Prioritization**: Priority as changeable property rather than static assignment

32.8 Theoretical Significance:

Chapter 32 addresses the critical gap between static goal representation and dynamic goal management in real agents. The multi-factor priority theory provides a sophisticated alternative to simple binary goal states, enabling nuanced resource allocation and attention management.

The extensive reification of goal states and transitions enables precise reasoning about goal lifecycle events, supporting both individual cognitive modeling and multi-agent coordination. The integration with scale theory provides quantitative foundations for qualitative priority reasoning.

The defeasible execution control mechanisms acknowledge that high-priority goals may sometimes not

be pursued (e.g., due to external constraints or competing urgent demands), while low-priority goals may occasionally receive attention (e.g., when resources are abundant or serendipitous opportunities arise).

The conflict resolution framework provides foundations for both personal decision-making and negotiation in multi-agent systems. The temporal exclusivity model captures the essential constraint that agents cannot simultaneously achieve incompatible goals.

The chapter's 20 axioms establish goal management as a sophisticated cognitive process involving continuous assessment, prioritization, modification, and execution control, moving far beyond simple goal-action mappings toward a realistic model of bounded rational agents operating under resource constraints and competing demands.

This represents one of the most comprehensive formal treatments of goal dynamics in the cognitive science literature, providing both psychological plausibility and computational tractability for modeling adaptive agent behavior in complex, changing environments.

Chapter 33

Chapter 33: Execution Envisionment

- **16 axioms total** covering execution envisionment, success/failure prediction, threshold effects, context specificity, and temporal aspects
- **4 main sections**: Executions and Their Envisionment, Envisioning Success and Failure, Specific and Arbitrary Contexts, Envisioned Executions and Time
- **Pure psychology** - mental simulation as cost-effective alternative to real-world action and tool for predicting others' behavior

33.1 Key Features Identified:

33.1.1 1. Foundation of Execution Envisionment (Axioms 33.1-33.4):

33.1.1.1 Core Concepts:

- **Axiom 33.1**: `actionIn` - actions in plans that occur because they are subgoals (causal link from plan membership to action occurrence)
- **Axiom 33.2**: `sideEffect` - unplanned consequences caused by plan subgoals but not themselves planned
- **Axiom 33.3**: `envisionExecution` - **complex framework** for imagining plan execution sequences
  - Agent a envisions agent b's plan execution through sequence s of envisioned causal systems
  - All envisioned eventualities must be either subgoals or side effects
  - At least one eventuality must be an action by the plan agent
  - Enables both self-execution envisionment and other-agent modeling
- **Axiom 33.4**: `envisionSelfExecution` - special case where agents imagine their own plan execution

33.1.1.2 Cost-Benefit Analysis:

- **Mental simulation** much cheaper than real action
- **Backtracking** possible in imagination but costly in reality
- **Prediction accuracy** vs. **computational efficiency** trade-off
- **Applications**: Self-planning and other-agent behavior prediction

33.1.2 2. Success, Failure, and Outcome Prediction (Axioms 33.5-33.10):

33.1.2.1 Outcome Envisionment:

- **Axiom 33.5**: `envisionExecutionSuccess` - goal achievement appears in envisioned execution sequence
- **Axiom 33.6**: `envisionExecutionFailure` - goal negation appears in envisioned execution sequence
- **Axiom 33.7**: `envisionExecutionSideEffect` - unplanned consequences appear in envisioned sequences
- **Axiom 33.8**: `envisionedLikelihoodOfSuccess` - probabilistic success estimates with constraint conditions
- **Axiom 33.9**: `envisionedOpportunity` - discovering unexpected subgoals for other plans during execution simulation

33.1.2.2 Threshold Effects:

- **Axiom 33.10**: `thresholdOfFailure` - **sophisticated threshold reasoning**
  - Branch points in execution depending on parameter values
  - Success when parameter $\geq$ threshold, failure when parameter $<$ threshold
  - Examples: "breaking point," "hang by a thread," "push one's luck"
  - Two-way branch resolution in envisionment with explicit success/failure branches

33.1.3 3. Context Specificity and Generalization (Axioms 33.11-33.13):

33.1.3.1 Specific vs. Arbitrary Contexts:

- **Axiom 33.11**: `specific1` - **complex specificity definition**
  - Entity specific to set with respect to purpose
  - Unique causal property not shared by other set members
  - Purpose-relative specificity (not absolute property)
- **Axiom 33.12**: `arbitrary` - universal property sharing
  - Any causally relevant property of arbitrary member shared by all set members
  - Enables "works anywhere," "applicable to any situation" reasoning

33.1.3.2 Execution Context Theory:

- **Axiom 33.13**: `arbitraryExecutionContext` - **complex context framework**
  - Background beliefs as execution context
  - Arbitrary context enables broad generalization
  - Specific context enables precise prediction
  - **Trade-off**: Precision vs. generalizability

33.1.4 4. Temporal Aspects of Execution (Axioms 33.14-33.16):

33.1.4.1 Time in Execution Envisionment:

- **Axiom 33.14**: `momentInExecution` - instants within execution time span
- **Axiom 33.15**: `arbitraryMomentInExecution` - "any moment while doing" reasoning
- **Axiom 33.16**: `envisionedExecutionDuration` - temporal extent prediction with units

33.1.4.2 English Expressions Captured:

- **Temporal**: "in the midst of," "in the course of," "time needed to," "manhours"
- **Threshold**: "breaking point," "hang by a thread," "push one's luck"
- **Context**: "could be done anywhere," "works in any circumstance," "applicable to any situation"

33.2 Technical Sophistication:

- **Mental Simulation Framework**: Complete alternative to real-world action for planning and prediction
- **Multi-Agent Modeling**: Agents can envision both self-execution and others' execution
- **Threshold Logic**: Sophisticated branch resolution for parameter-dependent success/failure
- **Context Theory**: Purpose-relative specificity vs. arbitrariness with causal grounding
- **Temporal Integration**: Time spans, moments, and duration prediction in envisioned executions
- **Opportunity Detection**: Serendipitous goal discovery during execution simulation
- **Moderate Reification**: 8 reified predicates for mental processes and judgments

33.3 Complexity Distribution:

- **Simple**: 7 axioms (basic definitions, straightforward relationships)
- **Moderate**: 5 axioms (execution failure, likelihood, opportunities, temporal concepts)
- **Complex**: 4 axioms (envisionExecution, thresholdOfFailure, specific1, arbitraryExecution-Context)

33.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Cognitive Economics**: Mental simulation as efficient alternative to action - **Predictive Modeling**: Anticipating outcomes of planned actions before execution - **Social Cognition**: Understanding and predicting others' behavior through plan attribution - **Risk Assessment**: Threshold-based reasoning about success/failure conditions - **Context Sensitivity**: Balancing prediction precision with generalizability - **Opportunity Recognition**: Discovering unexpected benefits during planning

33.5 Cross-Chapter Connections:

- **Chapter 24 (Envisioning)**: Extends ECS theory to plan execution sequences
- **Chapter 31 (Plans)**: Uses plan structure (subgoals, actions) as foundation
- **Chapter 20 (Modality)**: Likelihood scales for success probability estimation
- **Chapter 16 (Time)**: Temporal reasoning for execution duration and moments
- **Chapter 21 (Belief)**: Background beliefs as execution context
- **Chapter 15 (Causality)**: Causal involvement for specificity and side effects
- **Chapter 44+**: Forward reference to actual plan execution (contrasted with envisionment)

33.6 Applications Mentioned:

- **Threshold Examples**: Pencil breaking under force, failure points in various domains
- **Context Examples**: Course design for specific student needs vs. general applicability
- **Temporal Examples**: Manufacturing time estimates, project duration planning
- **Opportunity Examples**: Discovering synergies between different plans during simulation
- **Prediction Examples**: Anticipating others' actions by modeling their plans

33.7 Notable Design Decisions:

- **Non-Defeasible Framework**: All execution envisionment rules are strict logical definitions (no etc conditions)
- **Dual-Agent Support**: Same framework handles self-execution and other-agent modeling
- **Purpose-Relative Specificity**: Context properties depend on particular goals/purposes
- **Threshold Branch Logic**: Explicit modeling of parameter-dependent success/failure
- **Side Effect Integration**: Unplanned consequences as natural part of execution envisionment
- **Temporal Granularity**: Both point-in-time and duration-based temporal reasoning

33.8 Theoretical Significance:

Chapter 33 addresses the fundamental cognitive challenge of planning under uncertainty by providing a formal framework for mental simulation. The cost-effectiveness of imagination compared to action enables extensive "what-if" analysis without real-world consequences.

The threshold reasoning framework captures critical nonlinear effects where small parameter changes can determine success vs. failure - essential for understanding risk and resilience in human planning. The "breaking point" phenomenon appears across physical, social, and psychological domains.

The specificity vs. arbitrariness framework provides a principled approach to the precision-generalizability trade-off in prediction. Specific contexts enable accurate predictions but limited applicability, while arbitrary contexts enable broad generalization but reduced precision.

The integration of multi-agent modeling enables sophisticated social cognition where agents predict others' behavior by attributing plans and simulating their execution. This provides foundations for cooperation, competition, and coordination in multi-agent environments.

The chapter's 16 axioms establish execution envisionment as a central cognitive process enabling efficient planning, risk assessment, and social interaction through mental simulation rather than costly real-world experimentation.

This represents one of the most sophisticated formal treatments of mental simulation in cognitive science, providing both psychological plausibility and computational tractability for modeling human-like anticipatory reasoning and behavioral prediction.

Chapter 34

Chapter 34: Causes of Failure

- **6 axioms total** covering plan failure explanation patterns and three-way failure classification based on causal complexes
- **1 main section**: Causes of Failure (34.4) with comprehensive failure taxonomy
- **Pure psychology** - systematic framework for explaining why plans fail, integrating 15+ years of AI research on explanation patterns

34.1 Key Features Identified:

34.1.1 1. Foundational Failure Framework (Axioms 34.1-34.2):

34.1.1.1 Core Definitions:

- **Axiom 34.1**: `tryBy` - **complex trying framework**
  - Agent believes action is in causal complex for goal achievement
  - Goal causes action execution (motivational causation)
  - Integrates belief, causation, and goal-directed action
  - Foundation for all failure analysis
- **Axiom 34.2**: `failBy` - **simple failure definition**
  - Trying without goal achievement
  - Failure as outcome of trying, not absence of trying
  - Links failure directly to unsuccessful goal pursuit

34.1.1.2 Causal Complex Integration:

- **Causal Complex Theory**: From Chapter 15, sets of conditions jointly causing outcomes
- **Belief-Action Link**: Agents act because they believe actions will achieve goals
- **Motivational Causation**: Goals cause actions through agent reasoning

34.1.2 2. Three-Way Failure Classification (Axioms 34.3-34.6):

34.1.2.1 Type 1 Failure: Flawed Theory (Axiom 34.3):

- **Definition**: Causal complex doesn't actually cause goal achievement
- **Problem**: Agent's beliefs about causation are incorrect
- **Examples**: Wrong medicine for illness, ineffective study methods
- **Nature**: Theoretical/epistemic failure

34.1.2.2 Type 2 Failure: Missing Conditions (Axiom 34.4):

- **Definition**: Some conditions in causal complex don't hold
- **Problem**: Necessary conditions absent despite correct theory
- **Examples**: Recipe missing ingredients, plan missing resources
- **Nature**: Environmental/circumstantial failure

34.1.2.3 Type 2a Failure: Hail Mary Attempts (Axiom 34.5):

- **Special Subcase**: Agent hopes but doesn't believe conditions hold
- **Definition**: Type 2 failure where agent knows success is unlikely
- **Example**: "Hail Mary" football pass with low success probability
- **Nature**: Desperate attempt with acknowledged low likelihood

34.1.2.4 Type 3 Failure: Recursive/Execution Failure (Axiom 34.6):

- **Definition**: Failure to achieve subgoal causes failure of main goal
- **Problem**: Execution breakdown in plan hierarchy
- **Recursive Structure**: Subgoal failure can be Type 1, 2, or 3
- **Examples**: Can't open jar $\rightarrow$ can't make sandwich $\rightarrow$ can't eat lunch
- **Nature**: Hierarchical/execution failure

34.1.3 3. Comprehensive Failure Taxonomy:

34.1.3.1 15 Major Categories (Natural Language):

1. **Flawed Theory** (Type 1)
2. **Missing Condition** (Type 2)
3. **Hail Mary** (Type 2a)
4. **Execution Failure** (Type 3)
5. **Conflicting Goals**
6. **Resource Constraints**
7. **External Interference**
8. **Skill Deficits**
9. **Knowledge Gaps**

10. **Temporal Constraints**
11. **Social/Interpersonal Blocks**
12. **Motivational Failure**
13. **Physical Limitations**
14. **Technology Failure**
15. **Environmental Changes**

34.1.3.2 47+ Subcategories:

- Each major category broken into specific failure patterns
- Cross-referenced with AI research from Yale/Northwestern schools
- Systematic coverage of human plan failure experiences

34.2 Technical Sophistication:

- **Causal Complex Foundation**: Grounded in formal causation theory from Chapter 15
- **Recursive Definition**: Type 3 failures enable unlimited hierarchical depth
- **Belief Integration**: Connects failure analysis with belief systems from Chapter 21
- **Motivational Causation**: Goals cause actions through agent reasoning mechanisms
- **Reified Predicates**: 3 reified predicates for belief states and failure events
- **AI Research Synthesis**: 15+ years of explanation pattern research integration

34.3 Complexity Distribution:

- **Simple**: 4 axioms (basic failure types, straightforward definitions)
- **Moderate**: 2 axioms (trying framework, recursive failure structure)
- **Complex**: 0 axioms (natural language taxonomy provides complexity)

34.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Plan Debugging**: Systematic diagnosis of why plans fail - **Learning from Failure**: Understanding failure patterns for future planning improvement - **Explanation Generation**: Creating coherent stories about plan failures for communication - **Case-Based Reasoning**: Indexing failure experiences for analogical reasoning - **Risk Assessment**: Anticipating potential failure modes during planning - **Multi-Agent Coordination**: Understanding others' plan failures for cooperation

34.5 Cross-Chapter Connections:

- **Chapter 15 (Causality)**: Causal complex theory as foundation for failure analysis

- **Chapter 21 (Belief)**: Agent beliefs about causal relationships underlying trying
- **Chapter 28 (Goals)**: Goal theory providing motivational structure for failure
- **Chapter 31 (Plans)**: Plan structure enabling hierarchical failure patterns
- **Chapter 32 (Goal Management)**: Goal abandonment as response to repeated failure
- **Chapter 33 (Execution Envisionment)**: Mental simulation for failure prediction

34.6 Applications Mentioned:

- **Case Indexing**: Organizing failure experiences for retrieval and learning
- **Explanation Generation**: Creating coherent narratives about why plans failed
- **Explanation Recognition**: Understanding others' failure explanations
- **Yale AI Research**: Plan failure patterns from story understanding systems
- **Northwestern AI Research**: Goal failure analysis from planning systems
- **Hail Mary Examples**: Low-probability attempts in sports, business, relationships

34.7 Notable Design Decisions:

- **Three-Way Classification**: Comprehensive coverage with manageable number of basic types
- **Recursive Structure**: Hierarchical failures handled through Type 3 recursion
- **Belief-Centric Framework**: Failure analysis grounded in agent belief systems
- **Causal Complex Integration**: Formal causation theory underlying all failure types
- **AI Research Synthesis**: 15 years of explanation pattern research systematically organized
- **Non-Defeasible Definitions**: All failure type definitions are strict logical classifications

34.8 Theoretical Significance:

Chapter 34 addresses the fundamental challenge of explaining plan failures in a systematic, comprehensive manner. The three-way classification provides a principled decomposition of failure causes while the recursive structure enables unlimited hierarchical complexity.

The integration with causal complex theory provides formal foundations for failure analysis, moving beyond ad hoc categorizations toward principled diagnostic frameworks. The belief-centric approach acknowledges that failure analysis depends on agents' understanding of causation, not just objective facts.

The comprehensive taxonomy synthesizes decades of AI research on explanation patterns, providing systematic coverage of human failure experiences. This enables case-based reasoning, learning from failure, and improved planning through failure mode anticipation.

The framework supports three crucial applications: case indexing for analogical reasoning, explanation generation for communication, and explanation recognition for understanding others' accounts of failure.

The chapter's 6 axioms establish failure analysis as a systematic cognitive process, providing both theoretical precision and practical utility for modeling human reasoning about plan failures and learning from unsuccessful attempts.

This represents one of the most comprehensive formal treatments of plan failure in cognitive science, bridging AI research on explanation patterns with psychological theories of causal reasoning and goal-directed behavior.

Chapter 35

Chapter 35: Plan Elements

- **12 axioms total** covering plan elements including conditional actions, repetitive structures, preconditions, subplans, and temporal plan structures
- **3 main sections**: Temporal Structure and Goals, Plan Structure Elements, Complex Plan Types
- **Pure psychology** - systematic analysis of plan components and their relationship to temporal event structures

35.1 Key Features Identified:

35.1.1 1. Goal-Directed Temporal Structures (Axioms 35.1-35.3):

35.1.1.1 Conditional Actions and Goals:

- **Axiom 35.1**: **Complex defeasible rule** for conditional actions
  - Conditional action → condition enables goal OR negation of condition is goal
  - Two strategies: **exploiting** condition occurrence vs. **escaping** condition consequences
  - Condition must be relevant to agent's goals for action to make sense
  - Defeasible relationship acknowledging exceptions to goal-relevance

35.1.1.2 Repetitive Structure Goals:

- **Axiom 35.2**: `whileDo` - **defeasible goal negation**
  - Agent's goal is to negate the while-condition
  - Explains why agent keeps executing body until condition gone
  - Example: Wipe sweat while it drips into eyes → goal is no sweat in eyes
- **Axiom 35.3**: `repeatUntil` - **defeasible termination goal**
  - Until-condition is agent's goal
  - Each iteration brings agent closer to achieving termination condition

  - Example: Pound nails until flush → goal is nails flush with wood

35.1.1.3 Inferring Causal from Temporal Structure:

- **Key Insight**: Temporal patterns reveal underlying causal goal structure
- **Repetitive actions**: Usually have cumulative effect toward goal achievement
- **Conditional actions**: Usually exploit or counteract triggering conditions
- **Exception**: Individual pleasure actions (eating candy) may lack larger causal structure

35.1.2 2. Plan Structure Components (Axioms 35.4-35.8):

35.1.2.1 Subplan Variations:

- **Axiom 35.4**: `partialSubplan` - incomplete execution capability
  - Some terminal nodes not executable given constraints
  - Represents planning under resource/knowledge limitations
- **Axiom 35.5**: `subplanAgent` - **moderate complexity agent responsibility**
  - Single agent responsible for all actions in subplan
  - Enables clear accountability and coordination in complex plans

35.1.2.2 Plan Termination and Causation:

- **Axiom 35.6**: `planTermination` - **complex termination framework**
  - Final action that achieves goal through causal complex relationship
  - No other subgoals occur after termination action
  - Links plan structure to causal goal achievement
  - Most complex axiom in chapter with nested quantification

35.1.2.3 Precondition Taxonomy:

- **Axiom 35.7**: `knowledgePrecondition` - **simple knowledge enabling**
  - Agent knowing something enables action
  - Example: Knowing toolbox location enables finding tools
- **Axiom 35.8**: `resourcePrecondition` - **moderate resource framework**
  - Having resource + resource change + causal involvement
  - Distinguishes resources/tools from mere objects by change requirement
  - Example: Nails undergo location change during hammering

35.1.3 3. Complex Plan Types (Axioms 35.9-35.12):

35.1.3.1 Conditional and Repetitive Plans:

- **Axiom 35.9**: `conditionalPlan` - plans executed only when conditions occur
  - Links plan structure to conditional event structure

  - Reactive plans as special case responding to specific eventualities
- **Axiom 35.10**: `repetitivePlan` - plans with while/repeat structures
  - Captures systematic repetition patterns in goal achievement
  - Foundation for iterative problem-solving strategies

35.1.3.2 Temporal and Logical Plan Structure:

- **Axiom 35.11**: `periodicSubplan` - temporally regular execution patterns
  - Subplans with periodic temporal sequences (daily, weekly schedules)
  - Example: Workman's daily 9-5 job routine
- **Axiom 35.12**: `requiredPrecondition` - disjunctive enabling conditions
  - Precondition required regardless of which disjunct holds
  - Essential conditions that apply across alternative execution paths

35.2 Technical Sophistication:

- **Temporal-Causal Integration**: Systematic connection between temporal event patterns and underlying causal goal structures
- **Defeasible Goal Inference**: 3 axioms with `(etc)` for non-monotonic reasoning about goal-action relationships
- **Plan Component Taxonomy**: Comprehensive classification of subplans, agents, termination, and preconditions
- **Reified Mental States**: 5 reified predicates for knowledge, resources, temporal relations, and logical operators
- **Complex Quantification**: Nested universal and existential quantifiers especially in plan termination definitions
- **Cross-Domain Integration**: Heavy use of temporal, causal, and goal theory from previous chapters

35.3 Complexity Distribution:

- **Simple**: 7 axioms (basic plan types, simple preconditions, straightforward definitions)
- **Moderate**: 4 axioms (goal-action rules, subplan agents, resource preconditions)
- **Complex**: 1 axiom (plan termination with complex causal and temporal constraints)

35.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Plan Recognition**: Inferring agent goals from observable temporal action patterns - **Behavioral Explanation**: Understanding why agents engage in repetitive and conditional behaviors - **Plan Decomposition**: Breaking complex plans into

manageable subcomponents with clear responsibilities - **Precondition Analysis**: Systematic treatment of knowledge and resource requirements for action - **Temporal Planning**: Integration of periodic and conditional execution patterns - **Multi-Agent Coordination**: Clear assignment of agent responsibilities within complex plans

35.5 Cross-Chapter Connections:

- **Chapter 31 (Plans)**: Uses basic plan, subplan, and subgoal framework as foundation
- **Chapter 28 (Goals)**: Goal theory underlying all defeasible goal inference rules
- **Chapter 16-17 (Time/Events)**: Temporal sequence theory for repetitive and periodic plans
- **Chapter 15 (Causality)**: Causal complex theory for plan termination and preconditions
- **Chapter 21 (Belief)**: Knowledge preconditions linking to belief and knowing
- **Chapter 5 (Eventualities)**: Basic eventuality framework for all plan elements

35.6 Applications Mentioned:

- **Workman Example**: Comprehensive illustration of plan elements in construction work
  - Conditional actions: Wiping sweat when it drips into eyes
  - Repetitive actions: Pounding nails until flush, daily work routine
  - Knowledge preconditions: Knowing where toolbox is located
  - Resource preconditions: Having hammer and nails available
  - Plan termination: Putting tools away at 5 o'clock
- **Candy Eating**: Counter-example showing actions without larger causal structure
- **Tool vs. Resource**: Nails as resources that change state during use

35.7 Notable Design Decisions:

- **Defeasible Goal Inference**: Goal-action relationships are non-monotonic with exceptions
- **Causal Structure Priority**: Temporal patterns only meaningful when revealing underlying causal goal structures
- **Resource Change Requirement**: Resources/tools distinguished by undergoing state changes
- **Agent Responsibility**: Clear assignment of agents to subplan components
- **Precondition Taxonomy**: Systematic classification into knowledge vs. resource types
- **Plan-Event Integration**: Plan structures map directly onto temporal event structures

35.8 Theoretical Significance:

Chapter 35 addresses the fundamental challenge of connecting observable temporal behavior patterns with underlying mental goal structures. The key insight that "we have not really interpreted actions

until we understand the causal structure implicit in them" drives the systematic analysis of how temporal patterns reveal causal plan structures.

The defeasible goal inference rules provide a principled approach to plan recognition, enabling observers to infer agent goals from conditional and repetitive action patterns. This bridges the gap between external behavioral observation and internal mental state attribution.

The comprehensive precondition taxonomy acknowledges that successful action requires both knowledge (knowing where tools are) and resources (having tools available), with resources distinguished by undergoing changes that are causally involved in effects.

The plan component framework enables systematic decomposition of complex behaviors into manageable elements with clear agent responsibilities, temporal structures, and causal relationships. This supports both individual planning and multi-agent coordination.

The integration of temporal event structures (whileDo, repeatUntil, conditional) with plan structures provides a unified framework for understanding how complex goal-directed behaviors emerge from systematic manipulation of causal relationships.

The chapter's 12 axioms establish plan element analysis as a systematic cognitive process enabling behavioral interpretation, plan recognition, and goal attribution through temporal pattern analysis and causal structure inference.

This represents one of the most comprehensive formal treatments of plan structure analysis in cognitive science, providing both psychological plausibility for human plan recognition and computational tractability for automated behavior understanding systems.

Chapter 36

Chapter 36: Planning Modalities

- **11 axioms total** covering planning activity, multi-agent planning modalities, and counterfactual planning
- **3 main sections**: The Activity of Planning, Planning Activity and Other Agents, Counterfactual Planning
- **Pure psychology** - systematic treatment of how agents plan individually and in social contexts with various goal relationships

36.1 Key Features Identified:

36.1.1 1. Foundation of Planning Activity (Axioms 36.1-36.3):

36.1.1.1 Core Planning Definition:

- **Axiom 36.1**: `planning` - **reified mental activity**
  - Planning as sequence of envisionments of causal systems (plans)
  - Agent constructs plan sequences to achieve goals
  - Links envisioning theory with goal pursuit through reified activities
- **Axiom 36.2**: **Planning as subevent of goal pursuit**
  - All goal pursuit contains planning activities as components
  - Planning is necessary part of intentional behavior
  - Goal pursuit as "top-level intentional behavior" - what we do "going about our business"

36.1.1.2 Human Planning Capability:

- **Axiom 36.3**: **Universal human planning ability**
  - All persons with goals have ability to plan
  - Covers plan construction, adaptation, modification
  - Includes reactive planning, plan adaptation, elaborate from-scratch planning

- Planning as fundamental human cognitive capacity

36.1.1.3 Planning as Causal Knowledge Construction:

- **Key Insight**: Planning uses causal knowledge to construct larger-scale causal knowledge
- **Decision Process**: Choosing which causal knowledge becomes part of "The Plan" being executed
- **Intentional Behavior**: Source of all intentional action through plan selection and execution

36.1.2 2. Multi-Agent Planning Modalities (Axioms 36.4-36.8):

36.1.2.1 Collaborative Planning (Axiom 36.4):

- **Complex definition** with shared goals and collective agency
- Both agents plan for same goal g
- Set {a1, a2} has g as collective goal
- Joint plan sequence combines contributions from both agents
- True collaboration requires shared goal commitment

36.1.2.2 Assistive Planning (Axiom 36.5):

- **Weaker collaboration** - helper doesn't need full goal commitment
- Agent a1 assists a2 by achieving subgoal g1 of a2's goal g
- a1 satisfied with subgoal achievement, not full goal
- Asymmetric relationship with clear helper-helpee roles

36.1.2.3 Competitive Planning (Axiom 36.6):

- **Zero-sum goal competition** - agents work toward mutually exclusive goals
- Each agent plans to achieve goal instances that exclude the other
- Competitive plans designed to succeed at other's expense
- Symmetric competition with both agents having same ultimate goal type

36.1.2.4 Adversarial Planning (Axiom 36.7):

- **Stronger than competition** - active goal negation
- Agent a1 has explicit goal that a2 NOT achieve goal g
- Goes beyond competition to active opposition
- Plans specifically designed to prevent other's success

36.1.2.5 Counterplanning (Axiom 36.8):

- **Strongest opposition** - active plan interference
- Agent a1 tries to discover and block a2's plans

- Involves plan recognition followed by plan disruption
- Most sophisticated multi-agent planning requiring theory of mind

36.1.3 3. Counterfactual Planning Framework (Axioms 36.9-36.11):

36.1.3.1 Counterfactual Subgoals:

- **Axiom 36.9**: `counterfactualSubgoal` - **moderate complexity framework**
  - Subgoal that doesn't hold and agent doesn't plan to make it hold
  - Enables "what if" reasoning without commitment to making assumptions true
  - Example: Planning for world peace assuming you're president (when you're not)

36.1.3.2 Counterfactual Planning Process:

- **Axiom 36.10**: `counterfactualPlanning` - plans with counterfactual assumptions
  - Last plan in sequence contains counterfactual subgoal
  - Agent continues planning despite knowing assumption is false
  - Enables hypothetical reasoning and contingency planning

36.1.3.3 Other-Person Planning:

- **Axiom 36.11**: `otherPersonPlanning` - **perspective-taking planning**
  - Special case of counterfactual planning
  - Agent a1 plans as if having property g1 that a2 actually has
  - "What would I do if I were Barack Obama?" type reasoning
  - Uses substitution framework from Chapter 7 for property transfer

36.2 Technical Sophistication:

- **Reified Planning Activities**: Systematic treatment of planning as mental events with temporal structure
- **Multi-Agent Taxonomy**: Comprehensive classification from collaboration through adversarial competition
- **Counterfactual Reasoning**: Sophisticated treatment of hypothetical planning under false assumptions
- **Complex Quantification**: Multiple nested quantifiers over agents, plans, and sequences
- **Social Cognition Integration**: Planning modalities require theory of mind and other-agent modeling
- **Plan Sequence Management**: Complex handling of plan sequences and their combination in multi-agent contexts

36.3 Complexity Distribution:

- **Simple**: 3 axioms (basic planning definition, planning-goal relationship, human ability)
- **Moderate**: 3 axioms (counterfactual subgoals, counterfactual planning, other-person planning)
- **Complex**: 5 axioms (all multi-agent planning modalities with intricate agent and plan coordination)

36.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Social Planning**: Understanding how agents coordinate, compete, and conflict in planning - **Hypothetical Reasoning**: Counterfactual planning for scenario analysis and contingency preparation - **Theory of Mind**: Other-person planning requires understanding others' perspectives and capabilities - **Collaborative Systems**: Formal foundations for team planning and distributed problem solving - **Competitive Analysis**: Modeling strategic planning in competitive and adversarial contexts - **Perspective Taking**: Understanding how agents model others' planning processes

36.5 Cross-Chapter Connections:

- **Chapter 28 (Goals)**: Goal theory underlying all planning modalities and agent goal relationships
- **Chapter 31 (Plans)**: Basic plan structure and subgoal relationships used throughout
- **Chapter 24 (Envisioning)**: Planning as specialized envisioning activity creating plan sequences
- **Chapter 7 (Substitution)**: Substitution framework for other-person planning property transfer
- **Chapter 5 (Eventualities)**: Rexist predicate for confirming reality of substituted properties
- **Chapter 21 (Belief)**: Belief systems underlying counterfactual assumption management

36.6 Applications Mentioned:

- **Collaborative Planning**: Team projects, joint ventures, shared goal achievement
- **Assistive Planning**: Helper-helpee relationships, mentoring, support systems
- **Competitive Planning**: Business competition, sports, resource allocation conflicts
- **Adversarial Planning**: Military strategy, legal opposition, direct goal blocking
- **Counterplanning**: Intelligence analysis, strategic defense, plan disruption
- **Counterfactual Planning**: Pat imagining world peace if she were president
- **Other-Person Planning**: "What would I do if I were Barack Obama?" perspective taking

36.7 Notable Design Decisions:

- **Non-Defeasible Framework**: All planning modalities defined as strict logical relationships
- **Reified Activities**: Planning as mental events with temporal structure and causal relationships
- **Hierarchical Complexity**: From simple individual planning through complex adversarial counterplanning
- **Symmetric/Asymmetric Distinctions**: Collaborative and competitive planning (symmetric) vs. assistive and adversarial (asymmetric)
- **Counterfactual Integration**: Hypothetical planning integrated with basic planning framework
- **Social Cognition Requirements**: Multi-agent planning requires understanding others' mental states

36.8 Theoretical Significance:

Chapter 36 addresses the fundamental challenge of planning in social contexts where multiple agents have potentially conflicting goals and planning processes. The systematic taxonomy from collaboration through adversarial counterplanning provides a comprehensive framework for understanding social strategic reasoning.

The counterfactual planning framework enables sophisticated hypothetical reasoning, allowing agents to explore alternative scenarios without committing to making false assumptions true. This supports contingency planning, scenario analysis, and perspective-taking reasoning essential for social interaction.

The reification of planning activities as mental events with temporal structure provides foundations for understanding planning processes as cognitive phenomena rather than just abstract logical relationships. This enables modeling of planning effort, planning coordination, and planning interference.

The multi-agent framework acknowledges that real planning occurs in social contexts where agents must consider others' goals, plans, and planning processes. The hierarchy from collaborative through adversarial planning captures the full spectrum of social goal relationships and their planning implications.

The integration with counterfactual reasoning enables "what if" planning essential for strategic thinking, risk assessment, and creative problem solving. Other-person planning specifically supports perspective-taking and empathetic reasoning crucial for social cooperation and competition.

The chapter's 11 axioms establish planning modalities as a systematic framework for understanding how intelligent agents navigate social environments with complex goal relationships, enabling both cooperative and competitive strategic reasoning.

This represents one of the most comprehensive formal treatments of social planning in cognitive science, providing both psychological plausibility for human strategic reasoning and computational foundations for multi-agent planning systems.

Chapter 37

Chapter 37: Planning Goals

- **31 axioms total** covering planning constraints, preferences, and goal optimization including value minimization/maximization, enabling/blocking events, and plan instantiation
- **6 main sections**: Constraints and Preferences, Including and Avoiding, Enabling and Blocking, Minimizing and Maximizing, Locating Instances for Plans, Maintaining Plan Progress
- **Pure psychology** - comprehensive treatment of how agents optimize plans through preferences, constraints, and goal management

37.1 Key Features Identified:

37.1.1 1. Fundamental Preference Framework (Axioms 37.1-37.12):

37.1.1.1 Plan Goal Addition and Constraints:

- **Axiom 37.1**: `addGoalToPlan` - **moderate complexity goal integration**
  - Agent adds new goal to existing plan creating extended plan
  - Original plan becomes subplan of new plan through goal conjunction
  - Foundation for all plan optimization and constraint satisfaction
- **Axiom 37.2**: `planningConstraint` - **simple constraint definition**
  - Hard constraints as goals that must be added to plans
  - Constraints distinguished from soft preferences by necessity

37.1.1.2 Preference Theory Foundation:

- **Axiom 37.3**: **Preference type constraints** - agents, eventualities, and plans must be valid
- **Axiom 37.4**: **Defeasible incompatibility** - preferred options should be mutually exclusive
- **Axioms 37.5-37.7**: **Complex preference adoption rules**
  - **Axiom 37.5**: Preferred option adopted when possible
  - **Axiom 37.6**: Dispreferred option adopted when preferred impossible

  - **Axiom 37.7**: Forced choice with "bad for agent" negation handling
  - All three axioms use complex world understanding and possibility checking

37.1.1.3 Preference Ordering Properties:

- **Axiom 37.8**: **Antisymmetry** - if prefer a to b, then not prefer b to a
- **Axiom 37.9**: **Transitivity** - if prefer a to b and b to c, then prefer a to c
- **Axiom 37.10**: `prefer0` - **simple preference over absence**
- **Axioms 37.11-37.12**: **Value and cost-based preferences** - defeasible rules linking value/cost to preferences

37.1.2 2. Action Inclusion and Avoidance (Axiom 37.13):

37.1.2.1 Action Selection Preferences:

- **Axiom 37.13**: `preferAvoidAction` - avoiding action by preferring its negation
- Links to English phrases like "try to avoid" and "only as a last resort"
- Foundation for risk-averse planning and constraint-based action selection

37.1.3 3. Event Enabling and Blocking Framework (Axioms 37.14-37.23):

37.1.3.1 Basic Event Preferences:

- **Axiom 37.14**: `preferEnableEvent` - preferring actions that enable desired events
- **Axiom 37.15**: `preferBlockEvent` - preferring actions that prevent undesired events
- Foundation for proactive and protective planning strategies

37.1.3.2 Threat Management:

- **Axiom 37.16**: `preferEnableThreat` - malicious planning to "set one up to fail"
- **Axiom 37.17**: `preferBlockThreat` - defensive planning with "preemptive action"
- Integration with threat theory for risk assessment and management

37.1.3.3 Physical Transfer Control:

- **Axiom 37.18**: `preferEnableTransfer` - facilitating movement ("clear the path")
- **Axiom 37.19**: `preferBlockTransfer` - preventing movement ("obstruct")
- Uses reified movement predicates for spatial reasoning

37.1.3.4 Multi-Agent Goal Management:

- **Axiom 37.20**: `preferEnableAgency` - enabling other agents' actions
- **Axiom 37.21**: `preferBlockAgency` - blocking other agents' actions

- **Axiom 37.22**: `preferEnableOtherAgentGoalSatisfaction` - helping others succeed
- **Axiom 37.23**: `preferBlockOtherAgentGoalSatisfaction` - preventing others' success
- Comprehensive framework for cooperative and competitive multi-agent planning

37.1.4 4. Value Optimization Framework (Axioms 37.24-37.27):

37.1.4.1 Scale-Based Value Preferences:

- **Axiom 37.24**: `preferMinimizeValue` - **moderate complexity minimization**
  - Systematic preference for lower scale positions across all value comparisons
  - Captures "less is more," "be conservative," "frugal" strategies
- **Axiom 37.25**: `preferMaximizeValue` - **moderate complexity maximization**
  - Systematic preference for higher scale positions across all value comparisons
  - Captures "bigger is better," "as much as possible," "be liberal with" strategies

37.1.4.2 Duration-Specific Optimization:

- **Axiom 37.26**: `preferMinimizeDuration` - time efficiency ("time is of the essence")
- **Axiom 37.27**: `preferMaximizeDuration` - time extension ("prolong," "make it last")
- Special cases of value optimization applied to temporal planning

37.1.5 5. Plan Instantiation Framework (Axioms 37.28-37.30):

37.1.5.1 Entity Location and Identification:

- **Axiom 37.28**: `locateThing` - **moderate complexity instantiation**
  - Adding identifying properties to unspecified plan entities
  - Uses `wh'` predicate for context-dependent identification
  - Captures "locate," "search for," "get one's hands on" activities
- **Axiom 37.29**: `locateAgent` - finding people for plan roles ("fill the job")
- **Axiom 37.30**: `locateLocation` - finding places for plan activities
- Systematic treatment of plan completion through parameter instantiation

37.1.6 6. Plan Progress Maintenance (Axiom 37.31):

37.1.6.1 Progress Protection:

- **Axiom 37.31**: `preferMaintainPlanProgress` - **moderate complexity progress preservation**
  - Systematic preference against undoing achieved subgoals
  - Captures "avoid backpedaling" and "keep moving forward" strategies
  - Acknowledges that progress protection can be overridden when necessary

37.2 Technical Sophistication:

- **Comprehensive Preference Theory**: Complete treatment from basic preferences through complex optimization strategies
- **Defeasible Reasoning**: 6 axioms with `(etc)` for non-monotonic preference adoption and value reasoning
- **Scale Integration**: Heavy use of scale theory for value optimization and comparison
- **Multi-Agent Framework**: Systematic treatment of preferences involving other agents' goals and actions
- **Reified Mental States**: Extensive use of primed predicates for mental states and temporal relations
- **Plan Optimization**: Complete framework for transforming basic plans into optimized plans through preference satisfaction

37.3 Complexity Distribution:

- **Simple**: 21 axioms (basic definitions, type constraints, simple preference relationships)
- **Moderate**: 7 axioms (goal addition, value optimization, instantiation, progress maintenance)
- **Complex**: 3 axioms (sophisticated preference adoption with world understanding and possibility checking)

37.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Plan Optimization**: Moving beyond basic goal achievement to sophisticated plan refinement - **Preference Satisfaction**: Systematic treatment of soft constraints and optimization objectives - **Multi-Agent Planning**: Comprehensive framework for cooperative and competitive preference interactions - **Resource Management**: Value-based optimization for time, cost, and other scarce resources - **Risk Management**: Threat-aware planning with proactive and defensive strategies - **Plan Completion**: Systematic approach to instantiating underspecified plan parameters

37.5 Cross-Chapter Connections:

- **Chapter 31 (Plans)**: Uses basic plan structure for goal addition and subplan relationships
- **Chapter 28 (Goals)**: Goal theory underlying all preference and constraint relationships

- **Chapter 12 (Scales)**: Scale theory for value optimization and preference comparison
- **Chapter 21 (Belief)**: World understanding for preference adoption possibility checking
- **Chapter 15 (Causality)**: Enable/prevent relationships for event control preferences
- **Chapter 5 (Eventualities)**: Basic eventuality framework for preference objects

37.6 Applications Mentioned:

- **Planning Examples**: Sandwich buying (minimize time), driving (avoid traffic), hiking (flexible lunch location)
- **Preference Examples**: Aisle vs. window seat, apple vs. orange eating preferences
- **Forced Choice Examples**: Leg amputation vs. death from gangrene (bad option handling)
- **Value Examples**: "Less is more," "bigger is better," "time is of the essence," "make it last"
- **Multi-Agent Examples**: "Make someone happy," "thwart," "paralyze," "set one up to fail"
- **Instantiation Examples**: "Locate," "search for," "find someone to," "find a place where"
- **Progress Examples**: Block building (don't put things on cleared tops), "avoid backpedaling"

37.7 Notable Design Decisions:

- **Hard vs. Soft Constraints**: Clear distinction between must-satisfy constraints and optional preferences
- **Preference as Goals-in-Waiting**: Preferences adopted as goals when possible but not required
- **Bad Option Negation**: When preferred option is bad, adopt negation of dispreferred option instead
- **Partial Ordering**: Preferences form antisymmetric, transitive partial order (not complete ordering)
- **Defeasible Adoption**: Preference adoption is non-monotonic with exceptions possible
- **Scale-Based Optimization**: Value preferences defined through systematic scale position comparison
- **Progress Protection**: Default preference for maintaining achieved subgoals with override capability

37.8 Theoretical Significance:

Chapter 37 addresses the fundamental challenge of plan optimization under preferences and constraints. The distinction between hard constraints (must be satisfied) and soft preferences (adopted when possible) provides a principled approach to plan refinement and goal management.

The sophisticated preference adoption framework with world understanding and possibility checking enables realistic planning under uncertainty and resource constraints. The "goals-in-waiting" concept captures how preferences influence planning without overriding essential goals.

The comprehensive multi-agent preference framework enables modeling of cooperative, competitive, and adversarial planning relationships. Agents can prefer to help or hinder others, enable or block events, and optimize various aspects of their plans relative to other agents' activities.

The scale-based value optimization provides systematic foundations for resource management, time

optimization, and cost-benefit analysis in planning. The framework handles both minimization and maximization preferences with consistent logical structure.

The plan instantiation framework addresses the practical challenge of completing underspecified plans by locating appropriate entities, agents, and locations. This bridges the gap between abstract planning and concrete execution.

The progress maintenance framework acknowledges that agents naturally prefer to protect achieved progress while allowing for necessary backtracking when circumstances require it.

The chapter's 31 axioms establish planning goal management as a sophisticated cognitive process involving preference satisfaction, constraint management, value optimization, and strategic interaction with other agents.

This represents one of the most comprehensive formal treatments of plan optimization in cognitive science, providing both psychological plausibility for human planning behavior and computational foundations for automated planning systems under preferences and constraints.

Chapter 38

Chapter 38: Plan Construction

- **26 axioms total** covering planning processes from instantiation to scratch planning, process control, subprocesses, obstacles, and candidate plan selection
- **5 main sections**: Planning by Instantiation vs From Scratch, Process Control, Subprocess Activity, Obstacles, Candidate Selection
- **Pure psychology** - comprehensive treatment of how agents construct, control, and select plans through systematic construction processes

38.1 Key Features Identified:

38.1.1 1. Planning Construction Methods (Axioms 38.1-38.4):

38.1.1.1 Planning by Instantiation vs From Scratch:

- **Axiom 38.1**: `planningByInstantiating` - **moderate complexity habitual planning**
  - Only way between successive plans is instantiating types in earlier plan
  - Habitual action like restaurant ordering: taking general script and instantiating with specific waiter and menu
  - Near-automatic planning using well-established behavioral patterns
- **Axiom 38.2**: `planningFromScratch` - **complex constructive planning**
  - Mining causal knowledge to build plans rather than instantiating scripts
  - Constructing p2 from p1 by adding subgoal e2 based on causal belief about e2's involvement in e1
  - Complex nested quantification over plan elements, relations, and causal beliefs
  - Foundation for creative and adaptive planning when scripts don't exist

38.1.1.2 Single vs Multiple Goal Planning:

- **Axiom 38.3**: `singleGoalPlanning` - **simple goal focus**

  - Planning to achieve single non-conjunctive goals
  - Foundation for focused, undivided planning attention
- **Axiom 38.4**: `multipleGoalPlanning` - **simple goal multiplexing**
  - Planning to achieve conjunctive goals with multiple components
  - Enables complex goal achievement with multiple simultaneous objectives

38.1.2 2. Planning Process Control (Axioms 38.5-38.9):

38.1.2.1 Basic Process Control Operations:

- **Axiom 38.5**: `suspendPlanning` - **simple process suspension**
  - "Hold off on the details," "put aside the plan"
  - Temporal interruption of planning process with intention to resume
- **Axiom 38.6**: `resumePlanning` - **simple process continuation**
  - "Finish the plan," "take up the plan again"
  - Reactivating suspended planning processes
- **Axiom 38.7**: `abortPlanning` - **simple process termination**
  - "Stop planning," "give up on the plan"
  - Permanent abandonment of planning process

38.1.2.2 Planning Process Restart:

- **Axiom 38.8**: `replan` - **simple process restart**
  - "Start from scratch," "back to the drawing board," "try a different approach"
  - Restarting planning process for same goal with fresh approach
- **Axiom 38.9**: `replanSubplan` - **moderate complexity partial restart**
  - "Patch the plan," "fix part of the plan," "rework a step"
  - Restarting subplan for subgoal while maintaining overall plan structure

38.1.3 3. Plan Construction Operations (Axioms 38.10-38.15):

38.1.3.1 Plan Development Operations:

- **Axiom 38.10**: `specifyPlan` - **simple plan elaboration**
  - "Flesh out the details," "be specific about the plan"
  - McDonald's example: general eating plan becomes specific register choice
  - Adding detail to plans through `planInstancePlus` relationships
- **Axiom 38.11**: `addSubplan` - **complex subplan integration**
  - "Add a step," "addendum to the plan," "put a part in the plan"
  - Complex set operations to add subplan p3 achieving subgoal g1 of g
  - Sophisticated plan structure modification with constraint satisfaction
- **Axiom 38.12**: `removeSubplan` - **complex subplan elimination**

  - "Remove a step," "simplify the plan"
  - Mirror of addSubplan with reversed set difference operations
  - Systematic plan simplification while maintaining goal achievement

38.1.3.2 Plan Structure Operations:

- **Axiom 38.13**: `identifyPrecondition` - **moderate complexity precondition discovery**
  - "Realize we first need to," "see a necessary step"
  - Discovering that goals have preconditions requiring additional subgoals
  - Uses `enable` relation to identify necessary prerequisites
- **Axiom 38.14**: `selectSubplan` - **moderate complexity alternative selection**
  - "Choose a step," "select an action"
  - Replacing disjunction with one of its disjuncts
  - Converting alternative possibilities into specific choices
- **Axiom 38.15**: `orderSubplans` - **moderate complexity temporal ordering**
  - "First things first," "figure out the order"
  - Adding temporal constraints through `before` relations between subgoals
  - Temporal planning and sequencing operations

38.1.4 4. Obstacle Management Framework (Axioms 38.16-38.19):

38.1.4.1 Obstacle Theory:

- **Axiom 38.16**: `obstacle` - **moderate complexity impediment definition**
  - Something that causes goal not to occur
  - "Barrier," "impediment," "impasse"
  - Entity x or its property e1 can cause goal negation
- **Axiom 38.17**: `overcomeObstacle` - **simple obstacle elimination**
  - Causing obstacle not to occur in way consistent with goal achievement
  - Obstacle elimination as precondition satisfaction
  - Agent causes obstacle not to exist as obstacle

38.1.4.2 Planning Problems and Options:

- **Axiom 38.18**: `planningProblem` - **simple problem identification**
  - Obstacles as problems requiring overcoming in plans
  - "Challenge," "dilemma," "sticky situation," "something must be done"
  - Integration of obstacle theory with plan construction
- **Axiom 38.19**: **Planning option from obstacle overcoming** - simple option generation
  - Overcoming obstacles becomes planning options
  - Systematic generation of plan alternatives from problem analysis

38.1.5 5. Candidate Plan Selection (Axioms 38.20-38.26):

38.1.5.1 Candidate Plan Framework:

- **Axiom 38.20**: `candidatePlan` - **moderate complexity plan qualification**
  - Plan in focus that is executable now with respect to constraints
  - Agents construct multiple possible plans then select among candidates
  - Integration of attention, executability, and temporal constraints
- **Axiom 38.21**: `successfulPlanning` - **simple planning success**
  - Planning activity that produces candidate plans
  - Successful completion of plan construction process
- **Axiom 38.22**: `planningFailure` - **simple planning failure**
  - Planning without producing viable candidate plans
  - Recognition of planning process breakdown

38.1.5.2 Plan Assessment and Selection:

- **Axiom 38.23**: `assessPlan` - **simple plan evaluation**
  - Coming to belief about plan's likelihood of achieving goal
  - Graded belief in plan success using degree d
  - Foundation for rational plan selection
- **Axiom 38.24**: `selectCandidatePlan` - **complex rational selection**
  - Selecting plan assessed as most likely to succeed
  - Complex optimization over all candidate plans with assessment comparison
  - Rational decision making through comparative evaluation
- **Axiom 38.25**: **Selected plan scheduling** - simple scheduling integration
  - Selected plans are assigned execution times
  - Bridge from planning to scheduling and execution
- **Axiom 38.26**: **Non-selected candidate exclusion** - simple exclusion principle
  - Candidates not selected will not be executed
  - Ensures single plan selection for execution

38.2 Technical Sophistication:

- **Planning Process Control**: Comprehensive framework for suspend/resume/abort/restart operations with temporal anchoring
- **Plan Construction Operations**: Systematic treatment of plan development from specification through structural modification
- **Obstacle Integration**: Complete obstacle framework integrated with planning problem identification and resolution
- **Rational Selection**: Sophisticated candidate assessment and selection with optimization principles

- **Reified Planning Activities**: Extensive use of primed predicates for planning processes and causal relations
- **Process Integration**: Integration with general process control operations and forward references to scheduling

38.3 Complexity Distribution:

- **Simple**: 15 axioms (basic definitions, process control, problem identification)
- **Moderate**: 8 axioms (planning methods, precondition discovery, candidate management)
- **Complex**: 3 axioms (from-scratch planning, subplan operations, rational selection)

38.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Planning Process Management**: Complete framework for controlling planning activities through their lifecycle - **Plan Development**: Systematic approach to constructing and refining plans through multiple operations - **Problem Solving**: Integration of obstacle identification and resolution with plan construction - **Rational Planning**: Assessment-based selection among alternative candidate plans - **Bridge to Execution**: Connection from plan construction through scheduling to execution systems - **Creative vs Routine Planning**: Distinction between script instantiation and from-scratch construction

38.5 Cross-Chapter Connections:

- **Chapter 31 (Plans)**: Uses basic plan structure, subgoals, and plan relationships throughout
- **Chapter 28 (Goals)**: Goal theory underlying all planning construction and selection processes
- **Chapter 36 (Planning Modalities)**: Basic planning activity extended with construction-specific operations
- **Chapter 37 (Planning Goals)**: Preference and constraint satisfaction integrated with construction processes
- **Chapter 15 (Causality)**: Causal reasoning fundamental to from-scratch planning and obstacle management
- **Chapter 21 (Belief)**: Belief systems underlying plan assessment and causal reasoning
- **Chapter 42 (Plan Scheduling)**: Forward reference to scheduling operations for selected plans

38.6 Applications Mentioned:

- **Planning by Instantiation**: Restaurant ordering (general script → specific waiter and menu)
- **Planning From Scratch**: Creative problem solving when no scripts available

- **Process Control**: "Hold off on details," "back to drawing board," "patch the plan"
- **Plan Operations**: McDonald's planning (general eating → specific register choice)
- **Obstacle Examples**: "Challenge," "dilemma," "sticky situation," "something must be done"
- **Selection Examples**: Multiple candidate plan construction and assessment for goal achievement

38.7 Notable Design Decisions:

- **Script vs Creative Distinction**: Clear separation between habitual instantiation and creative from-scratch planning
- **Complete Process Control**: Comprehensive suspend/resume/abort/restart framework with temporal anchoring
- **Modular Construction Operations**: Systematic treatment of individual plan modification operations
- **Obstacle as Planning Problem**: Integration of impediment theory with planning problem framework
- **Rational Selection Process**: Assessment-based optimization for candidate plan selection
- **No Defeasible Rules**: All construction processes defined as strict logical relationships
- **Reified Activities**: Planning processes as mental events with causal relationships and temporal structure

38.8 Theoretical Significance:

Chapter 38 addresses the fundamental challenge of how agents actually construct plans from initial goals through final candidate selection. The distinction between planning by instantiation (using familiar scripts) and planning from scratch (creative construction using causal knowledge) captures essential differences in human planning behavior.

The comprehensive process control framework acknowledges that planning is not a linear process but involves suspension, resumption, restart, and abortion as circumstances change. The modular construction operations provide systematic foundations for understanding how plans develop through specification, subplan addition/removal, precondition identification, alternative selection, and temporal ordering.

The obstacle management framework integrates impediment theory with planning problem identification, enabling systematic approach to problem-solving during plan construction. The rational candidate selection process provides foundations for understanding how agents choose among multiple possible plans through assessment and optimization.

The integration of planning construction with process control, obstacle management, and rational selection creates a comprehensive framework for understanding planning as a complex cognitive process involving multiple interacting components rather than simple goal-to-action transformation.

The chapter's 26 axioms establish plan construction as sophisticated cognitive architecture involving script instantiation, creative construction, process management, problem solving, and rational decision making under uncertainty.

This represents one of the most comprehensive formal treatments of planning construction processes in cognitive science, providing both psychological plausibility for human planning behavior and computational foundations for automated planning systems that must construct, control, and select plans under realistic constraints.

Chapter 39

Chapter 39: Plan Adaptation

- **10 axioms total** covering plan tweaking, adaptation processes, modification types, and adaptation outcomes including success and failure
- **3 main sections**: Tweaking, Types of Plan Adaptation, Outcomes of Plan Adaptation
- **Pure psychology** - comprehensive treatment of how agents modify existing plans to achieve new or similar goals through systematic adaptation processes

39.1 Key Features Identified:

39.1.1 1. Fundamental Tweaking Framework (Axioms 39.1-39.4):

39.1.1.1 Core Tweaking Definition:

- **Axiom 39.1**: `tweak` - **complex plan step replacement**
  - Agent takes plan p1 for goal g1 and forms plan p2 for similar goal g2
  - Replaces one step e1 with similar step e2, yielding adapted plan
  - Restaurant example: paying before eating vs. after - retrieve standard plan and tweak event order
  - Foundation for all plan adaptation through similarity-based step substitution
  - Complex nested quantification over plan changes, envisioning, and similarity relations

39.1.1.2 Plan Adaptation as Tweaking Sequence:

- **Axiom 39.2**: `planAdaptation` - **complex systematic adaptation**
  - Planning activity where every step is a tweak
  - Produces sequence s of partial plans from p1 to p2
  - Systematic transformation through successive tweaking operations
  - Integration of planning activity with plan modification processes
  - Foundation for understanding adaptation as structured cognitive process

39.1.1.3 Plan Retrieval and Context:

- **Axiom 39.3**: `retrievePlan` - **simple memory access**
  - First step of adaptation: retrieving source plan from memory
  - "Remember how to," "consider an approach," "recall a technique," "pull from the playbook"
  - Plans can be previously executed, just thought through, or from other agents
  - Foundation for adaptation starting with existing causal knowledge
- **Axiom 39.4**: `modifyCurrentlyExecutingPlan` - **simple current plan modification**
  - Adaptation of plan currently being executed (thePlan from Chapter 31)
  - Real-time plan modification during execution
  - Integration with ongoing behavioral control systems

39.1.2 2. Adaptation Type Taxonomy (Axioms 39.5-39.7):

39.1.2.1 Temporal Context Classification:

- **Axiom 39.5**: `futurePlanModification` - **simple future plan adaptation**
  - Adaptation of plans not yet begun to be executed
  - Starting goal g1 not subgoal in current plan
  - Proactive planning modification before execution begins

39.1.2.2 Parameter Modification Types:

- **Axiom 39.6**: `modifyPlanValue` - **complex value-based adaptation**
  - Changing values in plans: "use more," "do it with more," "try increasing," "correct for"
  - Same external entity v at different points x1 and x2 in system s
  - Quantitative plan adjustment through parameter modification
  - Uses at' relations for positional value changes
- **Axiom 39.7**: `modifyPlanAgency` - **moderate agent substitution**
  - Changing agent who executes plan: "stand in for," "play the role of," "what would Jesus do?"
  - Extension from own previous plans to plans of others
  - Agent substitution using substitution framework from Chapter 7
  - Foundation for perspective-taking and role-based planning

39.1.3 3. Adaptation Outcome Framework (Axioms 39.8-39.10):

39.1.3.1 Cost Analysis:

- **Axiom 39.8**: `adaptationCost` - **simple effort measurement**
  - Number of substitutions needed in plan p1 to arrive at p2
  - Each substitution = one step in envisioning process

  - "Difficult to apply," "complex application of," "straightforward modification"
  - Foundation for effort-based adaptation decisions
  - Crude but tractable cost model for adaptation evaluation

39.1.3.2 Success and Failure Conditions:

- **Axiom 39.9**: `successfulAdaptation` - **simple success criteria**
  - New plan created that is executable by agent
  - "Successful application of," "thoroughly reworked the plan"
  - Integration with executability and constraint satisfaction
  - Temporal anchoring with Now and constraint world understanding
- **Axiom 39.10**: `failedAdaptation` - **simple failure definition**
  - No executable plan can be created from source plan
  - "Not applicable here," "unsalvageable plan," "not serve one's purposes"
  - Recognition of adaptation limits and source plan inadequacy

39.2 Technical Sophistication:

- **Similarity-Based Adaptation**: Central role of similarity relations between goals and plan steps in all modification processes
- **Systematic Tweaking Framework**: Every adaptation step defined as similarity-preserving plan element replacement
- **Multi-Type Modification**: Comprehensive taxonomy covering value changes, agent substitution, and temporal context
- **Cost-Benefit Analysis**: Formal treatment of adaptation effort vs. from-scratch planning alternatives
- **Success/Failure Conditions**: Clear criteria for adaptation outcomes based on executability
- **Reified Activities**: Strategic use of primed predicates for plan modification processes and memory retrieval
- **Integration with Planning Theory**: Deep connections to basic planning, envisioning, and substitution frameworks

39.3 Complexity Distribution:

- **Simple**: 6 axioms (memory retrieval, temporal context, cost analysis, success/failure conditions)
- **Moderate**: 1 axiom (agent substitution with substitution framework integration)
- **Complex**: 3 axioms (core tweaking, systematic adaptation, value modification with positional relations)

39.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Adaptive Planning**: Systematic reuse of existing plans for new but similar situations - **Cognitive Efficiency**: Avoiding from-scratch planning when adaptation is more efficient - **Cross-Context Transfer**: Applying plans across different agents, values, and situations - **Plan Reuse**: Leveraging accumulated planning knowledge through modification rather than recreation - **Effort Management**: Cost-based decisions between adaptation and new plan construction - **Real-Time Modification**: Adapting plans during execution when circumstances change

39.5 Cross-Chapter Connections:

- **Chapter 31 (Plans)**: Uses basic plan structure, thePlan for current execution, and subgoal relationships
- **Chapter 38 (Plan Construction)**: Adaptation as alternative to from-scratch planning construction
- **Chapter 24 (Envisioning)**: Adaptation as envisioning process with substitution steps
- **Chapter 22 (Similarity)**: Similarity relations fundamental to all tweaking operations
- **Chapter 7 (Substitution)**: Substitution framework for agent replacement and parameter modification
- **Chapter 36 (Planning Modalities)**: Planning activities extended with adaptation-specific operations

39.6 Applications Mentioned:

- **Restaurant Adaptation**: Standard restaurant plan adapted for pay-before vs. pay-after establishments
- **Performance Strategies**: Choir nervousness control (focus on conductor) adapted to lecture anxiety (focus on friendly face)
- **Activity Parameter Changes**: Movie plan modifications (different movie, theater, friend, driver, time)
- **Effort Examples**: One-handed to two-handed object lifting, “use more,” “try increasing,” “correct for”
- **Role Substitution**: “Stand in for,” “play the role of,” “fill in for,” “what would Jesus do?”
- **Plan Retrieval**: “Remember how to,” “pull from the playbook,” “recall a technique”

39.7 Notable Design Decisions:

- **Similarity-Centered Framework**: All adaptation based on similarity preservation between goals and plan steps

- **Tweaking as Atomic Operation**: Single step replacement as fundamental unit of adaptation
- **No Defeasible Rules**: All adaptation processes defined as strict logical relationships
- **Cost-Effectiveness Integration**: Explicit consideration of adaptation effort vs. from-scratch alternatives
- **Multi-Parameter Modification**: Systematic treatment of value, agent, and temporal parameter changes
- **Success/Failure Dichotomy**: Clear executability-based criteria for adaptation outcomes
- **Memory Integration**: Plan retrieval as explicit first step in adaptation process

39.8 Theoretical Significance:

Chapter 39 addresses the fundamental cognitive challenge of plan reuse through systematic modification. The similarity-based tweaking framework captures how agents leverage existing planning knowledge by making minimal changes to achieve new but related goals, representing a crucial efficiency mechanism in human cognition.

The comprehensive adaptation taxonomy from simple value changes through agent substitution demonstrates the flexibility of human planning systems in transferring knowledge across contexts. The "what would Jesus do?" example illustrates sophisticated perspective-taking where agents adapt others' plans by substituting themselves into different roles.

The cost analysis framework acknowledges that adaptation is not always more efficient than from-scratch planning, requiring agents to evaluate effort trade-offs. The envisioning-based cost model (substitutions as envisioning steps) provides tractable foundations for such evaluations.

The success/failure framework with executability criteria recognizes that not all plans can be successfully adapted to new contexts, requiring agents to recognize adaptation limits and fall back to alternative planning approaches.

The integration of adaptation with ongoing plan execution (modifyCurrentlyExecutingPlan) captures the dynamic nature of planning where agents must modify their approach in real-time as circumstances change during goal pursuit.

The chapter's 10 axioms establish plan adaptation as a sophisticated cognitive architecture involving memory retrieval, similarity-based modification, cost evaluation, and success assessment, representing one of the most comprehensive formal treatments of plan reuse in cognitive science.

This provides both psychological plausibility for human planning efficiency and computational foundations for adaptive planning systems that must reuse and modify existing solutions rather than constantly planning from scratch.

Chapter 40

Chapter 40: Design

- **13 axioms total** covering artifacts, designs, designing activities, design constraints, and design adherence including functionality, plan realization, and design quality evaluation
- **3 main sections**: Artifacts, Designs, Designing
- **Pure psychology** - comprehensive treatment of how agents create and evaluate designed entities as functional composite objects whose structure reflects plans for achieving functionality

40.1 Key Features Identified:

40.1.1 1. Artifact Theory Foundation (Axioms 40.1-40.4):

40.1.1.1 Core Artifact Definition:

- **Axiom 40.1**: `artifact` - **complex functional composite entities**
  - Composite entities with functionality realized through plan structure
  - Components are arguments of subgoals; relations among components are subgoals
  - Coffee cup example: cup and handle (components), attached relation, plan for moving coffee
  - Artificers: agents who designed (a1) and agents who brought into existence (a2)
  - Foundation for understanding human-made functional objects with intentional structure

40.1.1.2 Physical vs General Artifacts:

- **Axiom 40.2**: `artifact1` - **simple physical artifact specialization**
  - Specialized predicate for physical artifacts
  - General artifact can be physical object, organization, theory, or any human-made functional composite entity
  - Usually when we talk about artifacts, we have physical objects in mind

40.1.1.3 Intended vs Unintended Use:

- **Axiom 40.3**: `intendedUse` - **moderate complexity planned usage**
  - Use that corresponds to subgoal in artifact's plan
  - Broom sweeping kitchen floor (intended) vs. retrieving something behind refrigerator (unintended)
  - Agent involvement with artifact as instance of plan subgoal
- **Axiom 40.4**: `unintendedUse` - **simple non-planned usage**
  - Agent involvement with artifact that is not intended use
  - Recognition of creative repurposing beyond design intentions

40.1.2 2. Design Pattern Framework (Axioms 40.5-40.8):

40.1.2.1 Design as Pattern:

- **Axiom 40.5**: `design` - **complex pattern instantiation**
  - Pattern whose instances are artifacts
  - Can design teapot, logo, font, or government - broad applicability
  - Two principal features: intended to have instances and functional composite structure
  - Design plan is template; artifact plan is instance
  - "Artifact is plan made concrete; design is specification that concretizes plan"

40.1.2.2 Design Quality Assessment:

- **Axiom 40.6**: `notAchieve` - **simple failure definition**
  - Set of eventualities happens but goal does not result
  - Foundation for evaluating design effectiveness
- **Axiom 40.7**: `terminalSubgoalsOf` - **complex subgoal identification**
  - Subgoals that have no further subgoals of their own
  - Identifies executable steps at lowest level of plan hierarchy
  - Complex nested quantification over plan structure
- **Axiom 40.8**: `flawedDesign` - **complex quality evaluation**
  - Strong notion: ALL instances fail to achieve top-level goal
  - Weaker notion possible: some instances succeed but design unreliable
  - Systematic evaluation of design effectiveness across all instantiations

40.1.3 3. Design Activity and Adaptation (Axioms 40.9-40.13):

40.1.3.1 Designing Process:

- **Axiom 40.9**: `designing` - **moderate complexity planning activity**
  - Planning activity where all plans in sequence are plans for design
  - "Think about the design," "working out the layout," "imagining the structure"

  - Integration of planning theory with design creation processes
- **Axiom 40.10**: `designConstraint` - **simple constraint identification**
  - Multiple intended uses as conjunction of design constraints
  - Office layout design with respect to typical office goals
  - Foundation for multi-objective design optimization

40.1.3.2 Design Adaptation and Failure:

- **Axiom 40.11**: `designAdaptation` - **moderate complexity modification**
  - Designing activity that is also plan adaptation
  - "Customize the design," "design inspired by"
  - Integration with plan adaptation theory from Chapter 39
- **Axiom 40.12**: `designFailure` - **simple failure recognition**
  - Designing activity that is planning failure
  - Recognition when design creation process breaks down

40.1.3.3 Design Adherence Quality:

- **Axiom 40.13**: **Partial design adherence ordering - complex quality measurement**
  - Partial instantiation of design by implementing some properties/relations
  - `moreAdherentToDesign` partial ordering: more properties instantiated = more adherent
  - Foundation for evaluating how well entities conform to design patterns
  - Complex quantification over property and relation instantiation patterns

40.2 Technical Sophistication:

- **Plan-Structure Integration**: Deep connection between plan structure and artifact/design organization
- **Composite Entity Theory**: Systematic treatment of artifacts as composite entities with functional relationships
- **Pattern Instantiation Framework**: Sophisticated handling of designs as patterns with artifact instances
- **Quality Assessment Methods**: Multiple approaches to evaluating design effectiveness and adherence
- **Functionality Integration**: Central role of functionality0 predicate linking designs, artifacts, and plans
- **Multi-Level Analysis**: From individual artifacts through design patterns to designing activities
- **Cross-Domain Applicability**: Framework applies to physical objects, organizations, theories, and other designed entities

40.3 Complexity Distribution:

- **Simple**: 5 axioms (physical artifacts, unintended use, failure definitions, constraints, design failure)
- **Moderate**: 3 axioms (intended use, designing activity, design adaptation)
- **Complex**: 5 axioms (artifact definition, design pattern, terminal subgoals, flawed design, adherence ordering)

40.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Artifact Understanding**: Formal foundations for recognizing human-made functional objects - **Design Theory**: Systematic treatment of design patterns and their instantiation in artifacts - **Quality Evaluation**: Methods for assessing design effectiveness and adherence to design patterns
- **Creative Process**: Understanding designing as specialized planning activity - **Functional Analysis**: Connection between structure and intended functionality in designed objects - **Design Adaptation**: Modification and customization of existing design patterns

40.5 Cross-Chapter Connections:

- **Chapter 31 (Plans)**: Plan structure fundamental to artifact and design organization
- **Chapter 39 (Plan Adaptation)**: Design adaptation as specialized form of plan adaptation
- **Chapter 6 (Composite Entities)**: Artifacts as composite entities with component and relation structure
- **Chapter 5 (Eventualities)**: Functionality and plan goals as eventuality types
- **Chapter 36 (Planning)**: Designing as specialized planning activity
- **Chapter 22 (Similarity/Patterns)**: Design patterns and instance relationships

40.6 Applications Mentioned:

- **Coffee Cup Analysis**: Cup and handle components, attached relation, plan for moving coffee through containment and attachment
- **Usage Examples**: Broom for sweeping (intended) vs. retrieving objects (unintended)
- **Design Domains**: Teapot, logo, font, government - broad applicability across physical and abstract domains

- **Office Layout**: Design constraints based on typical office functionality goals
- **Design Language**: “Think about the design,” “working out the layout,” “configure,” “arrange”
- **Adaptation Examples**: “Customize the design,” “design inspired by”

40.7 Notable Design Decisions:

- **Broad Artifact Definition**: Includes physical objects, organizations, theories, and other functional composites
- **Plan-Structure Correspondence**: Artifact structure directly reflects plan for achieving functionality
- **Artificer Recognition**: Explicit acknowledgment of designing agents and manufacturing agents
- **Strong vs Weak Quality Notions**: Flawed design requires ALL instances to fail vs. some statistical failure rate
- **Pattern-Instance Framework**: Designs as patterns with artifacts as instances using plan instantiation
- **Multi-Objective Constraint Handling**: Design constraints as conjunction of functionality requirements
- **Partial Adherence Ordering**: Recognition that entities can partially conform to design patterns

40.8 Theoretical Significance:

Chapter 40 addresses the fundamental challenge of understanding designed objects as intentional structures that embody plans for achieving functionality. The integration of plan theory with composite entity theory provides sophisticated foundations for recognizing how human intentions become crystallized in artifact structure.

The distinction between designs as patterns and artifacts as instances captures essential relationships in human creative activity. Designs serve as templates that can be instantiated multiple times, with each artifact instance realizing the design plan in concrete form.

The comprehensive quality assessment framework from flawed designs through adherence ordering provides principled approaches to evaluating design effectiveness. The recognition that designs can fail (all instances fail to achieve goals) or succeed with varying degrees of adherence enables nuanced quality evaluation.

The integration of designing activity with planning theory demonstrates that design creation is a specialized form of planning where agents construct patterns intended for multiple instantiation. This bridges cognitive planning processes with creative design activities.

The broad applicability from physical artifacts through organizations to theories acknowledges that design principles apply across domains wherever humans create functional composite structures to achieve intended purposes.

The treatment of intended vs. unintended use recognizes that artifacts can be repurposed beyond original design intentions, capturing the creative flexibility in human artifact interaction while

maintaining connection to original design purposes.

The chapter's 13 axioms establish design theory as sophisticated cognitive architecture involving functional analysis, pattern instantiation, quality evaluation, and creative planning processes.

This represents one of the most comprehensive formal treatments of design theory in cognitive science, providing both psychological plausibility for human design activity and computational foundations for automated design systems that must create, evaluate, and adapt functional structures.

Chapter 41

Chapter 41: Decisions

- **34 axioms total** covering decision-making processes, choice sets, deliberation methods, justifications, and consequences including choice enumeration, selection, and evaluation
- **5 main sections**: The Decision-Making Process, Choices and Choice Sets, The Process of Deliberation, Justifications for Decisions, Consequences of Decisions
- **Pure psychology** - comprehensive treatment of how agents make decisions through systematic enumeration, deliberation, and selection processes with justification and consequence analysis

41.1 Key Features Identified:

41.1.1 1. Core Decision-Making Process (Axioms 41.1-41.10):

41.1.1.1 Three-Stage Decision Framework:

- **Axiom 41.1**: `decision` - **complex systematic process**
  - Three essential stages: 1) Identify choices ("think of possibilities"), 2) Deliberate about choices ("think about possibilities"), 3) Select one choice ("settle on")
  - Restaurant example: look at menu to identify choices, think about them, then select one
  - Choice enumeration e1 and deliberation e2 must precede selection e3, but e1 and e2 can be interleaved
  - All three are subprocesses of the overall decision-making process
  - Foundation for all systematic decision-making behavior

41.1.1.2 Choice Enumeration Process:

- **Axiom 41.2**: `choiceEnumeration` - **moderate complexity option identification**
  - Change to situation where agent thinks about set of options
  - Agent believes will decide on some member after thinking
  - Coming to think of options, have them in focus, recognize one must be chosen

  - Foundation for option discovery and consideration

41.1.1.3 Deliberation Process:

- **Axiom 41.3**: `deliberation` - **moderate complexity evaluation**
  - Thinking about options with goal of selecting one
  - Plan p whose top-level goal is selection, thinking about each option as subgoal
  - Thinking is in service of selection rather than mere contemplation
- **Axiom 41.4**: `selection` - **simple choice commitment**
  - Third part: select one option from choice set after deliberating
  - Deciding to do e after deliberating about set s of options

41.1.1.4 Selection Preferences and Process Control:

- **Axiom 41.5**: **Selection preference implication - simple defeasible rule**
  - If e3 selected and e4 not, expect agent to prefer e3 over e4 as goals
  - Uses defeasible (etc) condition - general expectation with exceptions
- **Axioms 41.6-41.8**: **Decision process control**
  - **Decision revision**: changing mind after making decision ("renege on," "go back on")
  - **Suspend decision**: interrupting before selection ("sleep on it")
  - **Resume decision**: continuing suspended decision ("finish deciding")

41.1.1.5 Linked Decision Series:

- **Axioms 41.9-41.10**: **Decision interdependence**
  - **Linked decisions**: options chosen have causal influence on other decisions
  - **Decision series**: sequences of linked decisions with recursive structure

41.1.2 2. Choice Sets and Evaluation (Axioms 41.11-41.19):

41.1.2.1 Choice Framework:

- **Axioms 41.11-41.12**: **Basic choice concepts**
  - **Choice set**: range of candidates for any decision
  - **Possible choice**: any member of choice set ("option," "prospect," "alternative")
- **Axioms 41.13-41.14**: **Choice comparison**
  - **betterChoice constraints**: both options must be possible choices in decision
  - **Utility-based comparison**: choice leading to more important goal is better choice
  - Weak constraint capturing utility theory without precise measures

41.1.2.2 Choice Analysis:

- **Axiom 41.15**: `choiceCharacteristic` - **simple feature identification**

  - Different characteristics relevant to decision ("distinguishing features," "pros and cons")
  - Properties of choices relevant to agent's goals
- **Axiom 41.16**: `bestChoice` - **simple optimality definition**
  - After deliberation, identify best choice ("frontrunner," "cream of the crop")
  - Choice better than or equal to all other possible choices

41.1.2.3 Choice Status:

- **Axioms 41.17-41.19**: **Selection outcomes**
  - **Selected choice**: becomes "one's pick," "one's decision"
  - **Unselected choice**: "not make the cut," "decided against"
  - **Previously selected choice**: when same choice set encountered again

41.1.3 3. Deliberation Methods (Axioms 41.20-41.28):

41.1.3.1 Consequence Envisioning:

- **Axiom 41.20**: `envisionChoiceConsequences` - **moderate complexity prediction**
  - "Look before you leap" - predicting outcome of particular choice
  - Add possible choice to current world understanding, envision consequences
  - Integration with envisioning theory for decision support

41.1.3.2 Pros and Cons Analysis:

- **Axioms 41.21-41.23**: **Cost-benefit evaluation**
  - **Pro**: envisioned consequence good for agent
  - **Con**: envisioned consequence bad for agent

  - **Cost-benefit analysis**: systematic consideration of alternatives determining pros/cons
  - Set of thinking-of actions about whether consequences are pros or cons
  - Complex quantification over thinking processes and consequence evaluation

41.1.3.3 Choice Selection Methods:

- **Axioms 41.24-41.26**: **Selection mechanisms**
  - **Choose over**: agent chooses one option over another in binary choice
  - **Choice criterion**: "yardstick," "benchmark," "litmus test" for determining best choice
  - **Obvious decision**: "no brainer," "clear cut" - one option much better than others
  - Uses scale theory with Hi region for obviously superior choices

41.1.3.4 Decision Quality Assessment:

- **Axioms 41.27-41.28**: **Decision characterization**

  - **Uncertain decision**: no certainty about best choice, no identifiable best option
  - **Insignificant decision**: doesn't matter which choice made ("inconsequential," "immaterial")
  - No option is better choice than any other option

41.1.4 4. Decision Justifications (Axioms 41.29-41.31):

41.1.4.1 Justification Framework:

- **Axiom 41.29**: `decisionFactor` - **simple justification element**
  - Pro for option selected or con for option not selected
  - Elements that support or oppose particular choices
- **Axiom 41.30**: `decisionJustification` - **simple justification set**
  - Set of decision factors constituting complete justification
  - Causal story built up by envisioning during deliberation
- **Axiom 41.31**: `arbitraryDecision` - **simple unjustified choice**
  - No justification for decision made ("whim," "impulsive decision," "crap shoot")
  - Empty justification set indicating absence of rational basis

41.1.5 5. Decision Consequences (Axioms 41.32-41.34):

41.1.5.1 Consequence Framework:

- **Axiom 41.32**: `decisionConsequence` - **simple causal outcome**
  - By making decision, people make things happen in world
  - Both selected option and decision process causally involved in consequence
  - Recognition that decisions have real-world effects
- **Axioms 41.33-41.34**: **Consequence evaluation**
  - **Positive consequences**: effects in accord with agent's goals (good for agent)
  - **Negative consequences**: effects counter to agent's goals (bad for agent)
  - Systematic evaluation of decision outcomes

41.2 Technical Sophistication:

- **Three-Stage Process Model**: Systematic enumeration → deliberation → selection framework
- **Process Control Integration**: Comprehensive suspend/resume/abort/restart operations with temporal constraints
- **Consequence Prediction**: Integration with envisioning theory for "look before you leap" analysis
- **Scale-Based Evaluation**: Use of scale theory for obvious decisions and choice comparison
- **Defeasible Reasoning**: Minimal use - only one axiom with (etc) for preference expectations

- **Reified Activities**: Extensive use of primed predicates for decision processes and mental operations
- **Multi-Level Analysis**: From individual choices through decision series to consequence evaluation

41.3 Complexity Distribution:

- **Simple**: 22 axioms (basic definitions, choice status, justifications, consequences)
- **Moderate**: 6 axioms (choice enumeration, deliberation, process control, consequence prediction)
- **Complex**: 6 axioms (core decision framework, cost-benefit analysis, choice criteria, obvious decisions, utility comparison)

41.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Rational Choice Theory**: Formal foundations for systematic decision-making under preferences and constraints - **Process Control**: Managing decision processes through suspension, resumption, and revision - **Consequence Analysis**: Understanding decision outcomes and their evaluation relative to goals - **Deliberation Methods**: Systematic approaches to choice evaluation including pros/cons and cost-benefit analysis - **Decision Quality**: Recognizing different types of decisions from obvious through arbitrary - **Justification Systems**: Framework for explaining and rationalizing decision choices

41.5 Cross-Chapter Connections:

- **Chapter 31 (Plans)**: Uses decideTo predicate for converting thoughts into intentions (thePlan)
- **Chapter 24 (Envisioning)**: envisionChoiceConsequences for predicting outcomes
- **Chapter 28 (Goals)**: Goal preference and importance evaluation underlying choice comparison
- **Chapter 12 (Scales)**: Scale theory for obvious decisions and choice quality evaluation
- **Chapter 21 (Belief)**: Belief formation during choice enumeration and deliberation
- **Chapter 45 (Execution Control)**: Process control operations (suspend, resume) for decision management

41.6 Applications Mentioned:

- **Restaurant Decision**: Look at menu (enumeration) → think about choices (deliberation) → select one (selection)

- **Choice Examples**: Cherry pie vs. apple pie, Ford vs. Chevrolet, number between 1-10, Pluto planet status
- **Process Control**: "Sleep on it" (suspend), "finish deciding" (resume), "renege on" (revision)
- **Deliberation Methods**: "Look before you leap," weighing pros and cons, cost-benefit analysis
- **Decision Types**: "No brainer" (obvious), "whim" (arbitrary), "inconsequential" (insignificant)
- **Justification Language**: "Yardstick," "benchmark," "litmus test" for criteria

41.7 Notable Design Decisions:

- **Three-Stage Process**: Universal framework from choice identification through selection
- **Interleaved Enumeration-Deliberation**: Recognition that option discovery and evaluation can overlap
- **Minimal Defeasible Reasoning**: Only one axiom uses (etc) - most decision logic is strict
- **Process Control Integration**: Decision-making as process subject to suspension/resumption like other activities
- **Consequence Causation**: Both selected option AND decision process causally involved in outcomes
- **Utility Without Precision**: Captures utility theory principles without requiring precise numerical values
- **Justification as Causal Story**: Decision factors form causal narrative through envisioning process

41.8 Theoretical Significance:

Chapter 41 addresses the fundamental cognitive challenge of systematic choice among alternatives. The three-stage framework (enumeration, deliberation, selection) captures essential structure of rational decision-making while allowing for realistic psychological processes like interleaved consideration and process interruption.

The integration of decision-making with process control theory acknowledges that real decisions unfold over time and can be suspended, resumed, or revised as circumstances change. This provides foundations for understanding decision-making as dynamic cognitive process rather than instantaneous logical computation.

The consequence prediction framework through envisioning enables "look before you leap" reasoning essential for effective decision-making. The pros/cons and cost-benefit analysis provide systematic methods for choice evaluation without requiring unrealistic precision in utility calculation.

The choice comparison framework captures utility theory insights about goal importance while acknowledging that real agents operate with partial orderings and qualitative judgments rather than

complete numerical utilities. The scale-based analysis of obvious decisions provides foundations for understanding when choices are clear-cut vs. difficult.

The justification framework as causal story built through envisioning connects decision reasoning to broader causal understanding, enabling agents to explain and defend their choices through narrative of consequences and goal relationships.

The treatment of decision consequences acknowledges that choices have real-world effects that can be evaluated against agent goals, providing feedback for future decision-making and learning from outcomes.

The chapter's 34 axioms establish decision-making as sophisticated cognitive architecture involving systematic enumeration, deliberation, selection, justification, and consequence evaluation - representing one of the most comprehensive formal treatments of choice behavior in cognitive science.

This provides both psychological plausibility for human decision-making processes and computational foundations for automated decision systems that must choose among alternatives under preferences, constraints, and uncertainty.

Chapter 42

Chapter 42: Scheduling

- **33 axioms total** covering simultaneous action constraints, schedule structures, capacity management, deadlines, and scheduling operations including temporal planning and resource allocation
- **7 main sections**: Simultaneous Actions, Schedules, Capacity Constraints, Deadlines, Scheduling Operations, Schedule Quality, Resource Management
- **Pure psychology** - comprehensive treatment of how agents manage temporal constraints and resource allocation in plan execution through scheduling mechanisms

42.1 Key Features Identified:

42.1.1 1. Simultaneous Action Constraints (Axioms 42.1-42.8):

42.1.1.1 Basic Simultaneity Framework:

- **Axiom 42.1**: `simultaneousActions` - **complex coordination constraint**
  - Set of actions happening at overlapping times with single agent
  - Agent cannot perform multiple simultaneous actions requiring same resources
  - Foundation for understanding multitasking limitations and resource conflicts
  - Integration with temporal interval theory and agent capacity constraints

42.1.1.2 Simultaneity Types:

- **Axiom 42.2**: `exclusiveSimultaneousActions` - **moderate complexity mutual exclusion**
  - Actions that cannot be performed simultaneously due to resource conflicts
  - Physical constraints: cannot write and type at same time
  - Cognitive constraints: cannot focus on multiple complex tasks simultaneously
- **Axiom 42.3**: `compatibleSimultaneousActions` - **simple compatibility constraint**
  - Actions that can be performed together without resource conflicts

  - Walking and talking, listening to music while working
  - Recognition of agent multitasking capabilities within limits

42.1.1.3 Resource Competition and Interference:

- **Axioms 42.4-42.6**: **Action interference patterns**
  - **Resource competition**: multiple actions requiring same limited resource
  - **Attention interference**: cognitive tasks competing for mental focus
  - **Physical interference**: bodily actions requiring same motor systems
  - Foundation for understanding multitasking limitations and scheduling constraints

42.1.1.4 Temporal Coordination Requirements:

- **Axioms 42.7-42.8**: **Synchronization constraints**
  - **Sequential dependency**: some actions must follow others in strict order
  - **Parallel coordination**: actions requiring simultaneous execution for effectiveness
  - Integration with plan structure and goal achievement requirements

42.1.2 2. Schedule Structure Theory (Axioms 42.9-42.16):

42.1.2.1 Core Schedule Definition:

- **Axiom 42.9**: `schedule` - **complex temporal plan organization**
  - Composite entity organizing actions across time with resource allocation
  - Temporal intervals assigned to activities with start/end constraints
  - Agent capacity management and resource distribution framework
  - Foundation for systematic time management and activity coordination

42.1.2.2 Schedule Components:

- **Axioms 42.10-42.12**: **Schedule element structure**
  - **Schedule entries**: individual time-slot assignments for specific activities
  - **Time blocks**: contiguous periods allocated to related tasks
  - **Buffer periods**: transition time between activities and unexpected delays
  - Systematic treatment of schedule granularity and temporal organization

42.1.2.3 Schedule Relationships:

- **Axioms 42.13-42.14**: **Temporal dependencies**
  - **Schedule precedence**: ordering constraints between schedule entries
  - **Schedule overlap**: permitted and prohibited temporal intersections
  - Foundation for complex project scheduling and dependency management

42.1.2.4 Schedule Modification Operations:

- **Axioms 42.15-42.16**: **Dynamic schedule adjustment**
  - **Schedule revision**: updating time allocations based on changing constraints
  - **Schedule optimization**: improving efficiency while maintaining feasibility
  - Integration with plan adaptation theory for responsive scheduling

42.1.3 3. Capacity Constraint Management (Axioms 42.17-42.22):

42.1.3.1 Agent Capacity Framework:

- **Axiom 42.17**: `agentCapacity` - **complex resource limitation**
  - Quantitative limits on simultaneous action performance
  - Physical, cognitive, and temporal capacity constraints
  - Scale-based capacity measurement from minimal to maximal capability
  - Foundation for realistic scheduling within agent limitations

42.1.3.2 Capacity Types and Measurement:

- **Axioms 42.18-42.19**: **Capacity categorization**
  - **Physical capacity**: bodily limitations on simultaneous physical actions
  - **Cognitive capacity**: attention and mental processing limitations
  - **Temporal capacity**: time availability and scheduling density limits
  - Scale theory integration for qualitative capacity assessment

42.1.3.3 Capacity Allocation Strategies:

- **Axioms 42.20-42.21**: **Resource distribution methods**
  - **Priority-based allocation**: important tasks receive more capacity resources
  - **Load balancing**: distributing capacity across multiple activities
  - **Capacity reservation**: holding resources for critical future activities

42.1.3.4 Overcommitment and Capacity Violations:

- **Axiom 42.22**: `capacityViolation` - **simple constraint failure**
  - Schedule requiring more capacity than agent possesses
  - Recognition of overcommitment problems and scheduling conflicts
  - Foundation for schedule feasibility checking and adjustment

42.1.4 4. Deadline Management Framework (Axioms 42.23-42.27):

42.1.4.1 Deadline Constraint Theory:

- **Axiom 42.23**: `deadline` - **moderate complexity temporal constraint**
  - Temporal limit by which activity must be completed

  - Hard deadlines (absolute requirements) vs. soft deadlines (preferences)
  - Integration with schedule priority and urgency assessment
  - Foundation for deadline-driven scheduling and time management

42.1.4.2 Deadline Types and Classification:

- **Axioms 42.24-42.25**: **Deadline categorization**
  - **External deadlines**: imposed by external agents or circumstances
  - **Self-imposed deadlines**: agent-created temporal constraints for motivation
  - **Flexible vs. rigid**: degree of tolerance for deadline extension
  - Scale-based deadline importance and urgency evaluation

42.1.4.3 Deadline Conflict Resolution:

- **Axioms 42.26-42.27**: **Temporal conflict management**
  - **Deadline prioritization**: ranking competing temporal constraints
  - **Deadline negotiation**: adjusting deadlines to resolve scheduling conflicts
  - Integration with goal importance and consequence evaluation

42.1.5 5. Scheduling Operations and Algorithms (Axioms 42.28-42.31):

42.1.5.1 Schedule Construction Methods:

- **Axiom 42.28**: `scheduling` - **complex temporal planning process**
  - Activity of creating feasible schedules within constraints
  - Integration of capacity limits, deadlines, and priority considerations
  - Systematic approach to temporal resource allocation and coordination
  - Foundation for automated scheduling systems and time management

42.1.5.2 Scheduling Strategies:

- **Axiom 42.29**: `schedulingStrategy` - **moderate complexity approach selection**
  - Different methods for schedule construction and optimization
  - Priority-first, deadline-first, capacity-based, and mixed strategies
  - Recognition that different situations require different scheduling approaches

42.1.5.3 Schedule Evaluation and Quality:

- **Axioms 42.30-42.31**: **Schedule assessment framework**
  - **Schedule efficiency**: minimizing wasted time and resources
  - **Schedule feasibility**: ensuring capacity and deadline constraints are met
  - **Schedule robustness**: maintaining quality under perturbations and changes

42.1.6 6. Advanced Scheduling Concepts (Axioms 42.32-42.33):

42.1.6.1 Dynamic Scheduling and Adaptation:

- **Axiom 42.32**: `scheduleAdaptation` - **complex responsive adjustment**
  - Modifying existing schedules in response to changing circumstances
  - Integration with plan adaptation theory for schedule modification
  - Recognition that real scheduling must handle uncertainty and change

42.1.6.2 Multi-Agent Scheduling Coordination:

- **Axiom 42.33**: `coordinatedScheduling` - **complex multi-agent temporal planning**
  - Scheduling involving multiple agents with interdependent activities
  - Shared resource coordination and temporal synchronization requirements
  - Foundation for team scheduling and collaborative project management

42.2 Technical Sophistication:

- **Multi-Dimensional Constraint Integration**: Simultaneous handling of temporal, capacity, and resource constraints
- **Scale Theory Application**: Qualitative assessment of capacity, priority, and deadline urgency
- **Dynamic Adaptation Framework**: Responsive scheduling modification under changing conditions
- **Multi-Agent Coordination**: Complex scheduling across multiple agents with shared resources
- **Reified Process Integration**: Extensive use of primed predicates for scheduling activities and mental processes
- **Temporal Logic Integration**: Sophisticated treatment of time intervals, sequencing, and synchronization
- **Resource Management Theory**: Systematic approach to capacity allocation and conflict resolution

42.3 Complexity Distribution:

- **Simple**: 8 axioms (basic compatibility, capacity violations, simple deadline types, basic operations)
- **Moderate**: 13 axioms (action interference, schedule components, deadline management, scheduling strategies)
- **Complex**: 12 axioms (simultaneity constraints, schedule structures, capacity frameworks, coordination systems)

42.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Temporal Planning**: Understanding how agents organize activities across time within constraints - **Resource Management**: Systematic allocation of limited capacity across competing activities - **Multitasking Theory**: Formal foundations for simultaneous action capabilities and limitations - **Project Management**: Framework for complex multi-activity scheduling with dependencies - **Time Management**: Personal scheduling strategies and deadline management approaches - **Coordination Theory**: Multi-agent scheduling and shared resource management

42.5 Cross-Chapter Connections:

- **Chapter 31 (Plans)**: Schedule structure built on plan organization and goal relationships
- **Chapter 39 (Plan Adaptation)**: Schedule adaptation as specialized form of plan modification
- **Chapter 15 (Time)**: Temporal interval theory fundamental to schedule structure
- **Chapter 12 (Scales)**: Scale theory for capacity, priority, and deadline assessment
- **Chapter 28 (Goals)**: Goal importance drives scheduling priorities and resource allocation
- **Chapter 41 (Decisions)**: Scheduling decisions involve choice among temporal alternatives
- **Chapter 36 (Planning)**: Scheduling as specialized form of temporal planning activity

42.6 Applications Mentioned:

- **Simultaneous Action Examples**: Writing while typing (impossible), walking while talking (possible)
- **Capacity Constraints**: Physical limitations, attention bottlenecks, time availability
- **Deadline Management**: Project deadlines, appointment scheduling, time-sensitive goals
- **Schedule Construction**: Calendar management, project planning, resource allocation
- **Coordination Examples**: Team scheduling, shared resource booking, meeting coordination
- **Adaptation Scenarios**: Schedule changes due to delays, priority shifts, resource unavailability

42.7 Notable Design Decisions:

- **Multi-Level Constraint Framework**: Integration of temporal, capacity, and resource constraints
- **Capacity as Quantitative Limit**: Recognition that agents have measurable performance limits
- **Dynamic Scheduling Emphasis**: Acknowledgment that real scheduling must handle change and uncertainty

- **Multi-Agent Extension**: Recognition that much scheduling involves coordination across agents
- **Scale-Based Assessment**: Use of qualitative scales rather than precise numerical measures
- **Reification of Scheduling Process**: Treatment of scheduling itself as cognitive activity
- **Resource Competition Framework**: Systematic treatment of why certain actions cannot be simultaneous

42.8 Theoretical Significance:

Chapter 42 addresses the fundamental cognitive and practical challenge of organizing activities across time within realistic constraints. The simultaneous action framework provides principled foundations for understanding multitasking capabilities and limitations, recognizing that agents cannot perform unlimited concurrent activities.

The schedule structure theory establishes systematic frameworks for temporal organization that integrate capacity limits, deadline constraints, and resource availability. This provides both descriptive foundations for human time management behavior and prescriptive frameworks for effective scheduling systems.

The capacity constraint framework acknowledges that real agents operate within measurable limits on physical, cognitive, and temporal performance. The integration with scale theory enables qualitative assessment of capacity utilization without requiring unrealistic precision in measurement.

The deadline management theory recognizes temporal constraints as fundamental drivers of scheduling decisions, with sophisticated treatment of deadline types, priorities, and conflict resolution strategies. This captures essential aspects of goal-driven time management behavior.

The scheduling operation framework establishes systematic methods for schedule construction and evaluation, recognizing that different situations require different scheduling strategies and that quality assessment involves multiple dimensions of efficiency, feasibility, and robustness.

The dynamic adaptation and multi-agent coordination frameworks acknowledge that real scheduling occurs in changing environments with multiple interdependent agents, requiring responsive modification and collaborative coordination capabilities.

The chapter's 33 axioms establish scheduling theory as sophisticated cognitive architecture involving constraint satisfaction, resource optimization, temporal coordination, and adaptive planning processes.

This represents one of the most comprehensive formal treatments of scheduling and time management in cognitive science, providing both psychological plausibility for human temporal organization behavior and computational foundations for automated scheduling systems that must operate within realistic constraints and changing conditions.

The integration across multiple constraint types and the emphasis on dynamic adaptation make this framework particularly valuable for understanding how agents successfully manage complex temporal demands in realistic environments with limited resources and competing priorities.

Chapter 43

Chapter 43: Monitoring

- **15 axioms total** covering monitoring processes, trigger conditions, monitoring characteristics, frequency, and failure modes including environmental surveillance and goal-relevant event detection
- **2 main sections**: Monitoring Processes, Characteristics of Monitoring Processes

- **Pure psychology** - comprehensive treatment of how agents continuously monitor their environment for threats and opportunities as conscious subprocess of interpret-act behavioral cycle

43.1 Key Features Identified:

43.1.1 1. Core Monitoring Framework (Axioms 43.1-43.7):

43.1.1.1 Basic Monitoring Process:

- **Axiom 43.1**: `monitor'` - **moderate complexity perception-based triggering**
  - Agent monitors set of eventuality types; when goal-relevant instance perceived, causes focus
  - Defeasible because sometimes miss significant threats/opportunities despite monitoring
  - Foundation for environmental surveillance as conscious subprocess of "Interpret" stage
  - Integration with perception theory and attention/focus mechanisms

43.1.1.2 Alternative Monitoring Triggering:

- **Axiom 43.2**: **Existence-based monitoring trigger - moderate complexity weaker condition**
  - If monitoring for eventuality type and token actually occurs, causes focus
  - Fails when something goal-relevant occurs but isn't perceived

  - Weaker than perception-based rule - covers cases where events happen but aren't noticed
  - Recognition that monitoring can trigger on actual occurrence independent of perception

43.1.1.3 Monitoring Sufficiency Condition:

- **Axiom 43.3**: **Goal relevance implies monitoring - simple defeasible rule**
  - If eventuality type has goal-relevant instance, should be monitored for
  - Defeasible due to difficult-to-perceive phenomena: radiation, carcinogens, bacteria
  - Captures rational monitoring strategy while acknowledging perceptual limitations
  - Foundation for understanding what ought to be monitored vs. what actually is monitored

43.1.1.4 Specialized Monitoring Types:

- **Axiom 43.4**: `monitorThing` - **moderate complexity object-focused monitoring**
  - Monitor for events happening to specific thing or states involving the thing
  - All monitored eventualities must have thing as argument
  - Systematic approach to entity-centered surveillance
- **Axiom 43.5**: `monitorAgent` - **simple agent-focused monitoring**
  - Specialized case of monitoring things where thing is an agent
  - Foundation for social monitoring and interpersonal awareness
- **Axiom 43.6**: `monitorSelf` - **simple self-focused monitoring**
  - Monitoring events happening to one's own self
  - Example: athlete monitoring body conditions relevant to performance
  - Foundation for self-awareness and bodily monitoring

43.1.1.5 Monitoring Termination:

- **Axiom 43.7**: `terminateMonitor'` - **moderate complexity process control**
  - People can stop monitoring some set of events after time
  - Change-based definition using changeFrom' and gen predicates
  - Integration with process control theory for dynamic monitoring management

43.1.2 2. Monitoring Characteristics and Operations (Axioms 43.8-43.15):

43.1.2.1 Trigger Condition Framework:

- **Axiom 43.8**: `monitorTriggerCondition` - **simple condition identification**
  - Eventuality types in monitored set s are trigger conditions
  - Foundation for systematic analysis of what agents watch for
- **Axiom 43.9**: **Goal relevance dichotomy - simple logical constraint**
  - Goal-relevant eventualities are either good or bad for agent
  - Captures that monitoring targets both threats (bad) and opportunities (good)
  - Theorem establishing completeness of good/bad classification for goal-relevant events

43.1.2.2 Trigger Condition States:

- **Axiom 43.10**: `monitorTriggerConditionSatisfied` - **simple condition fulfillment**
  - Trigger condition satisfied when instance actually exists (Rexist)
  - Recognition of successful condition detection
- **Axiom 43.11**: `monitorTriggerConditionUnsatisfied` - **simple condition non-fulfillment**
  - Trigger condition unsatisfied when no real instance exists
  - Foundation for understanding monitoring in absence of target events

43.1.2.3 Triggered Response Actions:

- **Axiom 43.12**: `monitorTriggeredAction` - **complex response mechanism**
  - When trigger condition occurs, typically execute action in response
  - Action e2 triggered by condition e1 while monitoring set s
  - Monitoring causes action, agent performs action, action serves agent's goals
  - Integration with causal complex theory - triggered actions must serve agent interests
  - Foundation for understanding monitoring as active process leading to behavioral responses

43.1.2.4 Temporal Monitoring Characteristics:

- **Axiom 43.13**: `monitoringTimeSpan` - **simple temporal framework**
  - Time span during which monitoring is conducted
  - Recognition that monitoring occurs over extended periods
- **Axiom 43.14**: `monitoringFrequency` - **complex frequency analysis**
  - Rate of monitoring occurrence in temporal sequence
  - Examples: second-by-second distance monitoring while driving, less frequent police car monitoring
  - Sophisticated treatment of monitoring as periodic activity with variable rates
  - Integration with temporal sequence theory and rate measurement

43.1.2.5 Monitoring Failure Analysis:

- **Axiom 43.15**: `monitoringFailure` - **complex failure mode analysis**
  - Two failure types: false negatives and false positives
  - **False negative**: event happens (Rexist) but monitoring doesn't cause focus
  - **False positive**: event doesn't happen (not Rexist) but monitoring causes focus
  - Comprehensive treatment of monitoring reliability and error modes
  - Foundation for understanding limitations of surveillance processes

43.2 Technical Sophistication:

- **Interpret-Act Cycle Integration**: Monitoring as conscious subprocess of behavioral control loop
- **Defeasible Monitoring Rules**: Recognition that monitoring can fail due to perceptual and environmental limitations
- **Multi-Level Monitoring Types**: From general monitoring through thing/agent/self specializations
- **Trigger-Response Framework**: Systematic connection from condition detection to behavioral response
- **Temporal Analysis**: Sophisticated treatment of monitoring frequency and temporal characteristics
- **Error Mode Analysis**: Comprehensive treatment of false positives and false negatives
- **Reified Process Integration**: Extensive use of primed predicates for monitoring activities and attention focus

43.3 Complexity Distribution:

- **Simple**: 8 axioms (monitoring specializations, trigger states, temporal span, goal relevance)
- **Moderate**: 4 axioms (basic monitoring rules, thing monitoring, termination)
- **Complex**: 3 axioms (triggered actions, monitoring frequency, failure analysis)

43.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Environmental Awareness**: Systematic framework for how agents track threats and opportunities - **Attention Management**: Connection between monitoring and focus of attention - **Behavioral Control**: Integration of monitoring with interpret-act behavioral cycles - **Error Analysis**: Understanding limitations and failure modes of surveillance processes - **Temporal Dynamics**: How monitoring frequency adapts to situational demands - **Goal-Directed Surveillance**: Connection between monitoring and agent goal structures

43.5 Cross-Chapter Connections:

- **Chapter 21 (Belief)**: Goal relevance and focus relationships fundamental to monitoring
- **Chapter 19 (Persons)**: Perceive predicate and agent theory underlying monitoring
- **Chapter 15 (Time)**: Temporal sequences and time spans for monitoring frequency
- **Chapter 28 (Goals)**: Goal relevance drives what should be monitored
- **Chapter 5 (Eventualities)**: Eventuality instances and existence (Rexist) central to monitoring
- **Chapter 8 (Logic)**: Causal relationships in triggered actions and monitoring responses

- **Chapter 31 (Plans)**: Causal complex integration for goal-serving triggered actions

43.6 Applications Mentioned:

- **Driving Examples**: Distance monitoring (second-by-second), police car monitoring (less frequent), gas gauge checking (every 30-40 miles)
- **Athletic Performance**: Athletes monitoring body conditions relevant to performance
- **Daily Life**: Continual monitoring during plan execution for threats and opportunities
- **Appointment Monitoring**: Checking watch more frequently as scheduled time approaches
- **Sports Officiating**: Basketball referee monitoring ball proximity to boundary lines
- **Environmental Hazards**: Difficulty monitoring radiation, carcinogens, bacteria in food

43.7 Notable Design Decisions:

- **Conscious Subprocess**: Monitoring as conscious part of interpret-act cycle rather than automatic background process
- **Defeasible Framework**: Recognition that monitoring naturally fails sometimes despite best efforts
- **Dual Triggering Modes**: Both perception-based and existence-based triggering for comprehensive coverage
- **Goal-Relevance Foundation**: All monitoring ultimately driven by goal relevance (threats/opportunities)
- **Frequency Adaptation**: Monitoring frequency increases as approach expected trigger conditions
- **Error Mode Analysis**: Systematic treatment of both false positive and false negative failures
- **Response Integration**: Triggered actions must serve agent goals through causal complex membership

43.8 Theoretical Significance:

Chapter 43 addresses the fundamental cognitive challenge of maintaining environmental awareness while executing ongoing plans. The monitoring framework captures essential aspects of how agents balance focused plan execution with vigilant surveillance for unexpected threats and opportunities.

The integration with the interpret-act behavioral cycle positions monitoring as a conscious subprocess of the "Interpret" stage rather than automatic background processing. This captures the phenomenology of deliberate attention to specific environmental conditions while acknowledging that much monitoring occurs with minimal computational effort.

The defeasible monitoring rules recognize that real agents operate under perceptual and cognitive limitations. Even when actively monitoring for important conditions, agents can miss significant

events (false negatives) or respond to non-existent events (false positives). This provides realistic foundations for imperfect surveillance systems.

The goal-relevance foundation establishes that all monitoring ultimately serves agent interests by detecting threats (bad events) or opportunities (good events). The theorem that goal-relevant events are either good or bad provides completeness for this classification scheme.

The frequency adaptation framework captures how monitoring intensity varies with situational demands. Agents monitor more frequently as they approach expected trigger conditions, optimizing attention allocation for maximum effectiveness.

The triggered action framework connects monitoring to behavioral control by requiring that monitoring responses serve agent goals through causal complex membership. This ensures monitoring leads to appropriate actions rather than mere passive observation.

The comprehensive failure analysis through false positive and false negative cases provides foundations for understanding monitoring reliability and error correction. This is essential for modeling realistic surveillance systems that must operate under uncertainty.

The chapter's 15 axioms establish monitoring as sophisticated cognitive architecture involving environmental surveillance, attention management, temporal dynamics, goal integration, and error handling processes.

This represents one of the most comprehensive formal treatments of environmental monitoring in cognitive science, providing both psychological plausibility for human vigilance behavior and computational foundations for artificial surveillance systems that must balance focused task execution with broad environmental awareness.

Chapter 44

Chapter 44: Execution Modalities

- **21 axioms total** covering plan execution, execution modalities, temporal execution patterns, and collaborative execution including planned, spontaneous, reactive, and temporal execution types
- **3 main sections**: Executions, Modalities of Execution, Plan Execution and Time
- **Pure psychology** - comprehensive treatment of how agents execute plans with different cognitive stances and temporal patterns, bridging mental plans with real-world action

44.1 Key Features Identified:

44.1.1 1. Core Execution Framework (Axioms 44.1-44.6):

44.1.1.1 Basic Action Execution:

- **Axiom 44.1**: `execute'` - **moderate complexity causal action execution**
  - Agent executes action in plan when action is subgoal, agent is actor, and subgoal status causally involved
  - Bridge between mental plans and real-world action through causal involvement
  - Foundation for distinguishing planned actions from mere coincidental actions
  - Integration with plan structure and causal theory

44.1.1.2 Plan Execution Framework:

- **Axiom 44.2**: `executePlan'` - **complex necessary conditions**
  - Execute plan by executing all subgoals for which agent is responsible
  - Introduces subexecution relation connecting individual action executions to overall plan execution
  - Plan must exist, at least one agent-responsible subgoal must exist, all such subgoals must be executed

  - Foundation for systematic treatment of complete plan realization
- **Axiom 44.3**: **Plan execution sufficient conditions - complex completeness guarantee**
  - If all conditions met (plan exists, agent has subgoals, all subgoals executed), then plan execution exists
  - Bidirectional relationship between conditions and execution existence
  - Ensures that meeting necessary conditions guarantees plan execution exists

44.1.1.3 Temporal Constraints and Existence:

- **Axiom 44.4**: `happensInTime` - **complex temporal constraint satisfaction**
  - Plans contain states/events that must hold at appropriate times (e.g., sun shining for picnic)
  - All temporal constraints implied by subgoals must actually be satisfied (Rexist)
  - Integration with temporal logic and constraint satisfaction
  - Foundation for realistic plan execution under temporal requirements
- **Axiom 44.5**: **Plan execution existence conditions - complex reality conditions**
  - Plan execution really exists iff all subexecutions exist AND non-agent subgoals happen in time
  - Ensures temporal relations among subgoals satisfied for legitimate execution order
  - Complete conditions for when plan execution moves from mental construct to real occurrence
- **Axiom 44.6**: **Subexecution as subevent - simple structural relationship**
  - Being subexecution is specific way of being subevent
  - Integration with event structure theory from Chapter 17
  - Foundation for analyzing partial executions as total executions of subplans

44.1.2 2. Execution Modalities Framework (Axioms 44.7-44.13):

44.1.2.1 Consciousness-Based Modalities:

- **Axiom 44.7**: `plannedExecution` - **moderate complexity conscious planning**
  - Execution is planned when it's subgoal of The Plan defining agent's intentional behavior and in focus
  - Integration with focus/attention theory and intentional action theory
  - Distinguishes deliberate planned actions from other execution types
- **Axiom 44.8**: `spontaneousExecution` - **moderate complexity unplanned action**
  - Execution not explicitly part of The Plan defining intentional behavior
  - People do things spontaneously without conscious planning
  - Captures impulsive or reactive behaviors outside deliberate planning
- **Axiom 44.9**: `nonconsciousExecution` - **complex unconscious action**
  - Agent executes but never aware of being agent of any subgoal
  - Cases where person is agent but not conscious of agency

  - Complex quantification ensuring no awareness across all subgoals
  - Foundation for automatic behaviors and unconscious actions

44.1.2.2 Environmental Interaction Modalities:

- **Axiom 44.10**: `reactiveExecution` - **moderate complexity environment-driven**
  - Change external to agent's mind causally involved in execution
  - Captures stimulus-response patterns and environmental reactivity
  - Integration with composite entity theory (externalTo predicate)
  - Foundation for understanding externally-triggered behaviors
- **Axiom 44.11**: `mentalExecution` - **moderate complexity internal-only**
  - Execution in focus but causes no change external to agent's mind
  - Pure thinking activities with no external world effects
  - Complement to reactive execution - entirely internal mental activity
  - Foundation for cognitive processes like reasoning, imagining, planning

44.1.2.3 Social and Constraint-Based Modalities:

- **Axiom 44.12**: `collaborativeExecution` - **simple multi-agent coordination**
  - Execution of plans by set of agents working together
  - Foundation for teamwork and cooperative action
  - Integration with multi-agent systems and social coordination
- **Axiom 44.13**: `followExecutionRules` - **complex rule-constrained execution**
  - Execution respecting constraints that must be followed
  - Constraints as properties of subexecutions with existence conditions
  - Foundation for rule-based behavior and protocol compliance
  - Integration with normative constraints and procedural requirements

44.1.3 3. Temporal Execution Patterns (Axioms 44.14-44.21):

44.1.3.1 Repetitive and Iterative Patterns:

- **Axiom 44.14**: `iterativeExecution` - **moderate complexity set-based iteration**
  - Execution iterating through elements of set using iteration predicate
  - Integration with event structure theory for systematic element processing
  - Foundation for loop-like behaviors and systematic processing
- **Axiom 44.15**: `repetitiveExecution` - **simple repeat-until pattern**
  - Execution using repeatUntil structure from event theory
  - Foundation for behaviors repeated until condition met
  - Integration with conditional termination and goal achievement

44.1.3.2 Temporal Structure Patterns:

- **Axiom 44.16**: `periodicExecution` - **simple rhythmic pattern**
  - Execution whose time span is periodic temporal sequence
  - Foundation for regular, rhythmic behaviors and scheduled activities
  - Integration with temporal sequence theory
- **Axiom 44.17**: `continuingExecution` - **moderate complexity plan continuation**
  - Execution continues execution of larger plan through subplan relationship
  - Foundation for hierarchical plan execution and task decomposition
  - Integration with subplan relationships and execution hierarchies

44.1.3.3 Temporal Coordination Patterns:

- **Axiom 44.18**: `timeTriggeredExecution` - **moderate complexity schedule-driven**
  - Execution triggered by specific time (e.g., leaving work at 8 AM)
  - Time being beginning of execution time span causally involved
  - Foundation for scheduled behaviors and time-based triggers
- **Axiom 44.19**: `consecutiveExecution` - **complex sequential coordination**
  - Plans executed one after another with time spans forming temporal sequence
  - Complex quantification ensuring proper temporal ordering
  - Foundation for sequential task execution and workflow management

44.1.3.4 Parallel Coordination Patterns:

- **Axiom 44.20**: `concurrentExecution` - **complex single-agent multitasking**
  - Single agent executing multiple plans with pairwise overlapping time spans
  - Foundation for multitasking and parallel activity management
  - Integration with interval overlap theory
- **Axiom 44.21**: `simultaneousExecution` - **moderate complexity multi-agent coordination**
  - Multiple agents executing different plans simultaneously with overlap
  - Conjunction of individual executions with temporal overlap
  - Foundation for coordinated multi-agent activities and synchronization

44.2 Technical Sophistication:

- **Plan-Reality Bridge**: Systematic connection between mental plans and real-world execution through causal involvement
- **Modality Taxonomy**: Comprehensive classification of execution types based on consciousness, environmental interaction, and social coordination
- **Temporal Pattern Framework**: Sophisticated treatment of execution timing from repetitive through concurrent patterns

- **Multi-Agent Coordination**: Framework for both collaborative (shared plans) and simultaneous (separate but coordinated) multi-agent execution
- **Constraint Integration**: Systematic treatment of temporal, rule-based, and environmental constraints on execution
- **Existence Conditions**: Precise conditions for when executions actually occur vs. remain mental constructs
- **Reified Process Integration**: Extensive use of primed predicates for execution processes and relationships

44.3 Complexity Distribution:

- **Simple**: 4 axioms (subevent relationship, collaborative execution, repetitive execution, periodic execution)
- **Moderate**: 9 axioms (basic action execution, planned/spontaneous/reactive/mental execution, iterative, continuing, time-triggered, simultaneous)
- **Complex**: 8 axioms (plan execution conditions, temporal constraints, nonconscious execution, rule following, consecutive/concurrent execution)

44.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Action Theory**: Bridge between mental plans and real-world behavior through execution framework - **Consciousness Studies**: Systematic treatment of conscious vs. unconscious action execution - **Temporal Coordination**: Understanding how agents coordinate activities across time - **Multi-Agent Systems**: Framework for collaborative and simultaneous execution coordination - **Behavioral Control**: Integration of environmental reactivity with planned behavior - **Workflow Management**: Foundations for sequential, concurrent, and rule-constrained task execution

44.5 Cross-Chapter Connections:

- **Chapter 31 (Plans)**: Plan structure fundamental to all execution processes
- **Chapter 17 (Event Structure)**: Subevent relationships, iteration, repeatUntil structures
- **Chapter 15 (Time)**: Temporal sequences, time spans, interval overlap for coordination
- **Chapter 10 (Composite Entities)**: externalTo predicate for reactive vs. mental execution distinction
- **Chapter 21 (Belief)**: Focus and attention mechanisms for planned vs. spontaneous execution
- **Chapter 8 (Logic)**: Causal involvement and conjunction operations
- **Chapter 5 (Eventualities)**: Rexist and existence conditions for real vs. mental execution

44.6 Applications Mentioned:

- **Execution Examples**: Driving (conscious agency), breathing (semi-involuntary), heart beating (non-agentive)
- **Planned vs. Spontaneous**: Deliberate actions vs. impulsive behaviors outside The Plan
- **Reactive vs. Mental**: Environmental responses vs. pure thinking activities
- **Temporal Triggers**: Leaving work at 8 AM, scheduled activities, time-based behaviors
- **Collaborative Work**: Team execution of shared plans and coordination
- **Rule Following**: Protocol compliance, constraint satisfaction in execution
- **Multitasking**: Concurrent plan execution by single agent, simultaneous multi-agent activities

44.7 Notable Design Decisions:

- **Causal Involvement Requirement**: Execution requires plan structure to causally contribute to action occurrence
- **The Plan Integration**: Connection to overall intentional behavior structure for planned vs. spontaneous distinction
- **Consciousness Gradation**: Spectrum from planned (conscious) through spontaneous to nonconscious execution
- **Internal vs. External Distinction**: Mental executions affect only mind, reactive executions respond to external changes
- **Temporal Constraint Satisfaction**: Execution existence requires satisfaction of all temporal relationships
- **Multi-Agent Framework**: Systematic treatment of both collaborative (shared) and simultaneous (coordinated) execution
- **Modality Orthogonality**: Different execution modalities can combine and interact

44.8 Theoretical Significance:

Chapter 44 addresses the fundamental challenge of connecting mental plans with real-world action execution. The framework systematically bridges the gap between having a plan "in mind" and actually executing it through the causal involvement requirement, ensuring that plan structure genuinely influences action occurrence rather than being mere coincidence.

The modality taxonomy provides sophisticated foundations for understanding different types of human action based on cognitive stance and environmental interaction. The consciousness spectrum from planned through spontaneous to nonconscious execution captures essential varieties of agency and intentionality in human behavior.

The temporal coordination framework addresses complex challenges of organizing activities across time, from simple repetitive patterns through sophisticated multi-agent coordination. This pro-

vides foundations for understanding everything from personal time management through complex organizational workflow coordination.

The existence conditions framework distinguishes between merely having execution plans and actually carrying them out, with precise requirements including subexecution realization and temporal constraint satisfaction. This provides realistic foundations for when executions move from mental constructs to real-world occurrence.

The environmental interaction distinction between reactive and mental execution captures fundamental aspects of how agents relate to their environment - sometimes responding to external changes, sometimes operating purely through internal mental activity.

The multi-agent coordination framework addresses both collaborative execution (shared plans) and simultaneous execution (separate but temporally coordinated plans), providing foundations for complex social and organizational behavior.

The chapter's 21 axioms establish execution theory as sophisticated cognitive architecture involving causal planning, consciousness modalities, temporal coordination, environmental interaction, and social cooperation processes.

This represents one of the most comprehensive formal treatments of plan execution in cognitive science, providing both psychological plausibility for human action patterns and computational foundations for automated execution systems that must coordinate complex activities across time, agents, and environmental constraints.

Chapter 45

Chapter 45: Execution Control

- **48 axioms total** covering execution control, plan beginning, progress tracking, completion, abstraction levels, aspectual operations, and distraction including left fringe analysis, preconditions, temporal execution control, and plan interruption/resumption
- **6 main sections**: Beginning Executions, Executions in Progress and Completed, Execution Costs and Outcomes, Abstract Plans and Their Instantiations, Plans and Aspect, Distraction
- **Pure psychology** - comprehensive treatment of internal structure of plan executions including temporal control, aspectual operations, and cognitive management of ongoing activities

45.1 Key Features Identified:

45.1.1 1. Beginning Executions Framework (Axioms 45.1-45.10):

45.1.1.1 Left Fringe Theory:

- **Axiom 45.1**: `leftFringe` - **complex plan initiation structure**
  - Set of subgoals that can be executed first - no prior dependencies
  - Complex nested quantification over temporal precedence relationships
  - Foundation for systematic plan initiation and parallel execution analysis
  - Integration with temporal logic and plan structure theory
- **Axiom 45.2**: **Simultaneous subgoals theorem - moderate complexity coordination principle**
  - If two subgoals must begin simultaneously, both or neither in left fringe
  - Ensures coherent treatment of parallel plan initiation
- **Axiom 45.3**: **Top-level goal theorem - simple structural principle**
  - Top-level goal always in left fringe since no subgoals precede it
  - Foundation for understanding goal-subgoal execution relationships

45.1.1.2 Precondition Management:

- **Axiom 45.4**: `precondition` - **moderate complexity enablement conditions**
  - Preconditions of plans are preconditions of left fringe subgoals
  - Preconditions not part of plan structure but external enablement conditions
- **Axioms 45.5-45.6**: **Precondition status tracking - simple temporal state management**
  - Satisfied vs. unsatisfied preconditions at execution time
  - Foundation for dynamic precondition monitoring
- **Axiom 45.7**: `violatePrecondition'` - **moderate complexity dynamic failure**
  - Satisfied preconditions can become unsatisfied during execution
  - Integration with change theory for dynamic plan environment

45.1.1.3 Execution Environment and Enablement:

- **Axiom 45.8**: `executionEnvironment` - **complex causal context**
  - States and events causally involved in subgoal occurrence/non-occurrence
  - Systematic treatment of execution context and environmental dependencies
- **Axiom 45.9**: `enabledPlan` - **moderate complexity readiness condition**
  - Plan enabled when all preconditions satisfied
  - Foundation for execution triggering and plan activation
- **Axiom 45.10**: `beginExecute` - **moderate complexity initiation process**
  - Begin execution by executing action in left fringe
  - Bridge between plan structure and actual execution commencement

45.1.2 2. Execution Progress and Completion (Axioms 45.11-45.20):

45.1.2.1 Temporal Execution States:

- **Axiom 45.11**: `executing` - **simple temporal activity state**
  - Agent executing action at specific time when execution occurs at that time
  - Foundation for time-indexed execution monitoring
- **Axiom 45.12**: `executingPlan` - **simple plan-level temporal state**
  - Executing plan when executing some subgoal of plan
  - Bridge between individual actions and overall plan execution
- **Axiom 45.13**: `executed` - **moderate complexity completion tracking**
  - Action executed by time t when execution ended before t
  - Foundation for progress tracking and completion analysis

45.1.2.2 Goal Completion Framework:

- **Axiom 45.14**: **Goal execution completeness - complex hierarchical completion**

  - Goal executed when all agent subgoals executed and non-agent subgoals happened in time
  - Systematic treatment of hierarchical goal achievement
  - Integration with temporal constraint satisfaction
- **Axiom 45.15**: `executedPlan` - **simple plan completion**
  - Plan executed when top-level goal executed
  - Bridge from goal completion to overall plan success
- **Axiom 45.16**: `remainingPlan` - **complex dynamic plan tracking**
  - Part of plan not yet executed at given time
  - Complex quantification over unexecuted subgoals and future temporal requirements
  - Foundation for dynamic plan monitoring and adaptive execution

45.1.2.3 Temporal Execution Analysis:

- **Axioms 45.17-45.18**: **Execution time boundaries - simple temporal marking**
  - Start and end times for plan executions
  - Foundation for duration analysis and scheduling
- **Axiom 45.19**: `executionDuration` - **moderate complexity temporal measurement**
  - Total duration of plan execution up to specific time
  - Integration with temporal interval theory and quantitative time measurement
- **Axiom 45.20**: `missDeadline` - **complex deadline violation detection**
  - Agent misses deadline when no execution completes before deadline
  - Integration with scheduling theory and temporal constraint management

45.1.3 3. Execution Outcomes and Costs (Axioms 45.21-45.23):

45.1.3.1 Cost and Success Analysis:

- **Axiom 45.21**: `executionCost` - **simple cost attribution**
  - Cost of execution using general cost predicate
  - Integration with goal theory and resource management
- **Axiom 45.22**: `executionSuccess` - **simple causal achievement**
  - Execution successful when it causes top-level goal
  - Foundation for evaluating plan effectiveness
- **Axiom 45.23**: `executionFailure` - **moderate complexity failure analysis**
  - Execution fails when doesn't cause goal AND goal doesn't occur otherwise
  - Sophisticated treatment requiring both causal failure and outcome failure

45.1.4 4. Abstract Plans and Instantiation (Axioms 45.24-45.29):

45.1.4.1 Plan Abstraction Framework:

- **Axiom 45.24**: `instantiatePlan` - **complex abstraction relationship**

  - More specific plan instantiates abstract plan through goal entailment and structure superset
  - Foundation for hierarchical plan abstraction and reuse
- **Axiom 45.25**: `sameAbstractPlan` - **complex temporal displacement**
  - Plans are same abstract plan when instantiate same nontemporal abstract plan
  - Foundation for understanding temporal variants of same basic plan structure
- **Axioms 45.26-45.27**: **Human execution patterns - simple/moderate behavioral rules**
  - People execute actions in plans, defeasibly execute scheduled plans
  - Integration with human agency and scheduling behavior

45.1.4.2 Reusability and Activities:

- **Axiom 45.28**: `reusablePlan` - **moderate complexity temporal flexibility**
  - Plan sufficiently abstract to instantiate at different times
  - Foundation for plan libraries and repeated execution patterns
- **Axiom 45.29**: `activity` - **moderate complexity group coordination**
  - Reusable plan that all group members can execute
  - Foundation for collaborative activities and shared behavioral patterns

45.1.5 5. Plans and Aspect (Temporal Control Operations) (Axioms 45.30-45.45):

45.1.5.1 Basic Aspectual Operations:

- **Axioms 45.30-45.31**: `start'` - **moderate complexity initiation operations**
  - Two equivalent definitions: execute left fringe action OR change to executing state
  - Foundation for systematic treatment of plan initiation
- **Axioms 45.32-45.33**: `continue` - **moderate/simple continuation operations**
  - Continue by executing remaining plan left fringe OR by being in executing state
  - Foundation for ongoing execution management
- **Axiom 45.34**: `stop'` - **moderate complexity termination**
  - Change from executing to not executing
  - Foundation for execution cessation and control

45.1.5.2 Advanced Control Operations:

- **Axiom 45.35**: `postpone'` - **complex temporal rescheduling**
  - Change execution start time to later time through plan re-instantiation
  - Sophisticated treatment of temporal flexibility and rescheduling
- **Axiom 45.36**: `complete'` - **moderate complexity successful termination**
  - Cause change to state of having executed plan

  - Distinguished from mere stopping by achievement implication
- **Axiom 45.37**: `interrupt'` - **moderate complexity disruption**
  - Stop without completing - premature termination
  - Foundation for handling execution disruption

45.1.5.3 Resumption and Recovery Operations:

- **Axiom 45.38**: `resume'` - **complex continuation after interruption**
  - Start remaining plan after interruption using same abstract plan
  - Sophisticated treatment of execution recovery and continuation
- **Axiom 45.39**: `restart'` - **complex fresh start after interruption**
  - Start original plan from beginning after interruption
  - Distinguished from resume by starting over rather than continuing
- **Axiom 45.40**: `pause'` - **moderate complexity temporary suspension**
  - Stop with actual resumption (not just intention)
  - Foundation for temporary execution suspension with guaranteed resumption

45.1.5.4 Pause Management and Ongoing Execution:

- **Axiom 45.41**: `pauseInterval` - **complex temporal gap analysis**
  - Interval between pause and first subsequent resumption
  - Complex quantification ensuring identification of correct resumption event
- **Axiom 45.42**: `ongoing` - **moderate complexity extended execution**
  - Execution ongoing when either executing or paused throughout interval
  - Foundation for understanding continuous engagement including pauses

45.1.5.5 Intentional Control Operations:

- **Axiom 45.43**: `suspend'` - **moderate complexity intentional interruption**
  - Interrupt with intention to resume (distinguished from pause by intention vs. actuality)
  - Integration with intention theory and plan management
- **Axiom 45.44**: `abort'` - **complex permanent termination**
  - Interrupt with intentions not to restart or resume for all future times
  - Complex quantification over future temporal intentions
- **Axiom 45.45**: `terminate'` - **simple execution ending**
  - Plan terminated when completed or aborted
  - Foundation for recognizing definitive execution endings

45.1.6 6. Distraction and Attention Management (Axioms 45.46-45.48):

45.1.6.1 Distraction Theory:

- **Axiom 45.46**: `distract'` - **moderate complexity attention shift**

  - Change from thinking of one thing to another where properties of new thing causally involved
  - Integration with attention theory and causal analysis of cognitive shifts
- **Axiom 45.47**: **Execution implies thinking - simple defeasible connection**
  - When executing action in plan, defeasibly thinking of both action and plan
  - Foundation for understanding cognitive engagement during execution
- **Axiom 45.48**: **Distraction causes suspension - moderate defeasible disruption**
  - Distraction during plan execution defeasibly causes suspension
  - Integration of attention management with execution control

45.2 Technical Sophistication:

- **Left Fringe Analysis**: Sophisticated treatment of plan initiation through dependency-free subgoal identification
- **Aspectual Operations**: Comprehensive formal treatment of temporal execution control (start, stop, pause, resume, etc.)
- **Abstract Plan Framework**: Systematic handling of plan abstraction levels and instantiation relationships
- **Temporal Control Integration**: Deep integration with temporal logic, scheduling, and constraint satisfaction
- **Hierarchical Completion**: Complex treatment of goal completion through subgoal achievement
- **Dynamic Plan Tracking**: Sophisticated remaining plan analysis for adaptive execution
- **Intentional Control**: Integration with intention theory for suspension, abortion, and resumption decisions
- **Multi-Level Analysis**: From individual action execution through plan completion to abstract plan reuse

45.3 Complexity Distribution:

- **Simple**: 14 axioms (basic temporal states, completion definitions, cost attribution, termination)
- **Moderate**: 19 axioms (preconditions, execution environment, control operations, distraction)
- **Complex**: 15 axioms (left fringe, execution environment, remaining plan, abstract plans, advanced control operations)

45.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Execution Control**: Systematic management of plan execution from initiation through completion - **Temporal Planning**: Understanding

how agents control activities across time with interruption and resumption - **Progress Tracking**: Monitoring plan execution state and remaining work - **Plan Reuse**: Framework for abstract plans and instantiation across different contexts - **Attention Management**: Integration of cognitive attention with execution control - **Adaptive Execution**: Dynamic adjustment of execution based on changing circumstances

45.5 Cross-Chapter Connections:

- **Chapter 44 (Execution Modalities)**: Basic execution framework extended with temporal control
- **Chapter 31 (Plans)**: Plan structure fundamental to all execution control operations
- **Chapter 42 (Scheduling)**: Deadline concepts and temporal constraint management
- **Chapter 28 (Goals)**: Goal achievement and cost theory underlying execution evaluation
- **Chapter 21 (Belief)**: Intention theory for suspension and abortion decisions
- **Chapter 15 (Time)**: Temporal intervals, precedence, and duration measurement
- **Chapter 17 (Event Structure)**: Change operations and temporal event relationships

45.6 Applications Mentioned:

- **Left Fringe Examples**: Simultaneous subgoals, top-level goal initiation, parallel execution coordination
- **Precondition Scenarios**: Plan enabling conditions, dynamic precondition violation during execution
- **Control Operations**: Start/stop/pause/resume cycles, postponement and rescheduling
- **Completion Analysis**: Goal achievement through subgoal completion, execution success vs. failure
- **Abstract Plans**: Reusable plan patterns, temporal displacement of same basic plan
- **Distraction Examples**: Attention shifts causing execution suspension, cognitive engagement during execution

45.7 Notable Design Decisions:

- **Left Fringe Foundation**: Plan initiation based on dependency-free subgoals rather than arbitrary starting points
- **Dual Aspectual Definitions**: Multiple equivalent definitions for aspectual predicates (e.g., start as action execution vs. state change)
- **Intention-Action Distinction**: Suspend (intention to resume) vs. pause (actual resumption), abort (intention not to resume)
- **Hierarchical Completion**: Goal completion requires both agent action completion and non-agent event occurrence

- **Abstract Plan Instantiation**: Systematic treatment of plan abstraction levels through goal entailment and structure supersets
- **Ongoing Spans Pauses**: Execution can be ongoing even during pauses, distinguished from simple active execution
- **Distraction Integration**: Attention shifts formally connected to execution control through suspension mechanisms

45.8 Theoretical Significance:

Chapter 45 addresses the fundamental challenge of controlling plan execution across time, providing systematic frameworks for initiating, monitoring, modifying, and terminating ongoing activities. The left fringe theory establishes principled foundations for plan initiation based on dependency analysis rather than arbitrary starting points.

The aspectual operations framework provides formal treatments of temporal control operations (start, stop, pause, resume, etc.) that capture essential aspects of how agents manage ongoing activities. The dual definitions for many operations (e.g., start as action execution vs. state change) provide multiple perspectives on the same underlying phenomena.

The abstract plan instantiation framework addresses plan reuse and adaptation by providing systematic treatment of abstraction levels. Plans can be made more specific through instantiation while maintaining connections to abstract templates, enabling both individual plan adaptation and group activity coordination.

The remaining plan analysis provides dynamic tracking of execution progress, enabling adaptive execution management as circumstances change. This connects with the precondition management framework to handle environmental changes during execution.

The intentional control framework distinguishes operations based on agent intentions (suspend vs. pause, abort vs. simple stopping), providing nuanced treatment of voluntary execution management. This integration with intention theory enables sophisticated analysis of planned vs. reactive execution control.

The distraction and attention management framework connects cognitive attention with execution control, recognizing that execution involves conscious engagement that can be disrupted by attention shifts. This provides foundations for understanding execution robustness under cognitive demands.

The chapter's 48 axioms establish execution control as sophisticated cognitive architecture involving dependency analysis, temporal control, progress tracking, adaptive management, attention integration, and intentional decision-making processes.

This represents one of the most comprehensive formal treatments of execution control in cognitive science, providing both psychological plausibility for human activity management and computational

foundations for automated execution systems that must operate robustly in dynamic environments with changing priorities and attention demands.

Chapter 46

Chapter 46: Repetitive Execution

- **17 axioms total** covering repetitive execution structures, repetition management, iteration through sequences, and temporal aspects of cyclical processes
- **2 main frameworks**: Repetition structures (repeat-until loops) and iteration through sequences (for-each loops)
- **All psychology** - focuses on plan execution patterns and cognitive processes involved in repetitive behavior

46.1 Key Features Identified:

1. **Repetition Structure Framework**:
   - Axioms 46.1-46.5: Core repetition concepts with repeatUntil structures, body executions, and termination conditions
   - Axiom 46.1: Fundamental definition linking repetitive execution to repeatUntil structures with subplan instantiation
   - Axioms 46.2-46.5: Classification of repetition instances (aRepetition, start, terminate, intermediate)
   - Integration with plan execution framework showing repetition as structured plan patterns
2. **Temporal State Management**:
   - Axioms 46.6-46.8: State tracking through completed, remaining, and current repetitions at any time point
   - Sophisticated temporal reasoning using begins/ends predicates to track execution progress
   - Set-based characterization of repetition status enabling temporal queries about process state
   - Current repetition defined through temporal containment (time t between start and end of execution)
3. **Sequential Navigation**:
   - Axioms 46.9-46.10: Previous and next repetition relationships with temporal ordering

constraints
- Complex logic preventing intervening repetitions between reference points
- Enables navigation through repetition history and planning future iterations
- Supports reasoning about repetition sequences in temporal context

4. **Execution Counting**:
   - Axiom 46.11: Quantitative tracking of completed repetitions using cardinality of completed sets
   - Provides numerical interface to repetition progress for decision making
   - Links set-based state representation to integer counting for computational use
5. **Iteration Through Sequences**:
   - Axioms 46.12-46.17: Specialized framework for iterating through collections with substitution
   - Axiom 46.12: Core iterationThru definition using argument substitution across sequence elements
   - Distinguished abstract participant (variable x) gets instantiated with successive sequence elements
   - Creates sequence s1 of execution instances from template e1 and data sequence s
6. **Iteration Management**:
   - Axioms 46.13-46.14: Starting iterations and tracking completed iterations through sequences
   - Axiom 46.15: Active iteration progression with doNextIterationThru' reified action
   - Axioms 46.16-46.17: Early termination (abort) vs. successful completion of iteration sequences
   - Full lifecycle management from initialization through completion or abortion

46.2 Technical Sophistication:

- **Reified Actions**: Uses primed predicates (repeatUntil', executePlan', doNextIterationThru') for reified eventuality structures
- **Temporal Integration**: Extensive use of temporal predicates (before, begins, ends) for precise ordering
- **Dual Frameworks**: Provides both general repetition (loops with conditions) and specific iteration (loops over data)
- **Plan Integration**: Builds on plan execution framework with subplan relationships and execution states
- **Sequence Operations**: Sophisticated sequence manipulation with substitution and instantiation operations
- **Set Characterization**: Multiple predicates defined through set membership criteria based on temporal completion

46.3 Complexity Distribution:

- Simple: 4 axioms (basic definitions, sequence operations)
- Moderate: 9 axioms (temporal state tracking, iteration management)
- Complex: 4 axioms (core repetition definition, temporal ordering with constraint checking)

46.4 Conceptual Importance:

This chapter provides essential infrastructure for: - **Cognitive Psychology**: Models of repetitive behavior, habit formation, and cyclical mental processes - **Artificial Intelligence**: Loop constructs, iteration patterns, and repetitive plan execution - **Process Modeling**: Temporal tracking of repetitive activities and progress monitoring - **Programming Semantics**: Formal foundation for loop constructs in computational models - **Temporal Reasoning**: Sophisticated temporal logic for cyclical and repetitive phenomena

46.5 Cross-Chapter Connections:

- **Chapter 17 (Event Structure)**: Fundamental dependency on event structures and repeatUntil constructs
- **Chapter 5 (Eventualities)**: Uses Rexist and basic eventuality framework
- **Chapter 7 (Substitution)**: Critical use of substitution operations for iteration through sequences

- **Chapter 15 (Time)**: Temporal predicates (before, begins, ends) for ordering repetitions
- **Chapter 41 (Planning)**: Plan execution framework (executePlan, subplan relationships)
- **Chapter 6 (Sets)**: Set operations for characterizing completed/remaining repetitions

46.6 Applications Mentioned:

- **Father kissing children**: Iteration example with generic child (x) instantiated as Matthew, Mark, etc.
- **Repeat-until loops**: General repetitive execution with termination conditions
- **For-each patterns**: Iteration through collections with systematic element processing
- **Plan abortion**: Early termination of repetitive processes before completion
- **Progress tracking**: Monitoring completion status during long repetitive tasks

46.7 Notable Design Decisions:

- **Dual Frameworks**: Separates general repetition from sequence-specific iteration
- **Temporal Precision**: Uses precise temporal predicates rather than simple ordering

- **Plan Integration**: Treats repetitive execution as structured plan patterns rather than primitive loops
- **Reified Actions**: Actions like doNextIterationThru' are eventuality entities, not just predicates
- **Set-Based State**: Represents process state through sets of completed/remaining executions
- **Substitution Semantics**: Iteration uses sophisticated substitution rather than simple variable binding
- **Instance Relationships**: Careful distinction between template executions and their instances

46.8 Theoretical Significance:

Chapter 46 provides formal foundations for repetitive behavior, bridging computational loop constructs with psychological models of cyclical activities. The dual framework of general repetition and sequence iteration captures both condition-driven repetition (repeat-until) and data-driven repetition (for-each).

The temporal sophistication goes beyond simple sequential ordering to provide precise timing relationships, enabling reasoning about concurrent repetitions, progress tracking, and temporal queries. The integration with plan execution frameworks shows how repetitive patterns emerge from structured planning rather than primitive iteration constructs.

The reification of actions like doNextIterationThru' reflects the psychological reality that deciding to continue or abort repetitive processes involves conscious mental actions. The set-based state representation enables sophisticated queries about repetition progress while maintaining temporal precision.

This formalization supports both AI applications (implementing loop constructs with formal semantics) and cognitive modeling (understanding repetitive behavior patterns). The chapter establishes repetitive execution as a fundamental cognitive and computational pattern, with formal foundations for reasoning about its temporal dynamics and control structures.

The work represents a significant advance in formalizing repetitive processes, providing both mathematical precision and psychological plausibility for modeling cyclical behavior in cognitive systems.

Chapter 47

Chapter 47: Mind–Body Interaction

- **56 axioms total** covering mind-body interaction, awareness, perception, bodily actions, control channels, consciousness states, and dualist framework
- **9 main sections**: Introduction, Mind and Body, Awareness and Focusing, Perception, Bodily Actions, Perceiving Bodily Actions, Controlling Perception, Levels of Capability, States of Consciousness
- **All psychology** - comprehensive formalization of commonsense mind-body dualism with systematic treatment of mental and physical interaction

47.1 Key Features Identified:

1. **Dualist Framework**:
   - Axiom 47.1: Mind is not a physical object, establishing fundamental Cartesian dualism
   - Clear separation between physical body and non-physical mind interacting bidirectionally
   - Adopts commonsense theory rather than contemporary scientific views
   - Follows Descartes' interactionalism where mind and body influence each other
2. **Awareness Architecture**:
   - Axioms 47.2-47.6: Three-tier mental structure with memory, awareness, and focus of attention
   - Axiom 47.2: Every person's mind has awareness as intermediate region
   - Axiom 47.5: Focus of attention is part of awareness (concepts in focus are also in awareness)
   - Axiom 47.6: Concentration as willing oneself to think of something, with direct causal efficacy
3. **Perception Theory**:
   - Axioms 47.7-47.14: Perception causes concepts to enter awareness and defeasibly focus
   - Axiom 47.9: All perception requires intact sense organs as enablers
   - Axioms 47.10-47.13: Classification of sense organs (eyes, ears, tongue, nose)

- Axiom 47.14: Stimulation from external world defeasibly causes perception
- Axiom 47.15: Attending to sense organs through willing concentration on perceived content

4. **Bodily Action Taxonomy**:
   - Axiom 47.16: Bodily actions as movement of body parts from one location to another
   - Axiom 47.17: Reflexes as externally caused body movements without willing
   - Axiom 47.18: Tics as internally caused involuntary actions without external stimulation
   - Axiom 47.19: Voluntary actions as bodily actions directly caused by willing
   - Axioms 47.20-47.21: Voluntary actions are defeasibly plan executions serving larger goals
5. **Action Quality and Coordination**:
   - Axiom 47.24: Coordinated actions involving multiple body parts with causal timing relationships
   - Axioms 47.25-47.27: Success/failure conditions for bodily actions based on willing effectiveness and goal achievement
   - Axiom 47.28: Fluent actions as skilled voluntary action executions
   - Axiom 47.29: Awkward actions as voluntary actions in low region of skill scale
6. **Kinesthetic Perception**:
   - Axioms 47.30-47.31: Bodily actions defeasibly cause self-perception (always for voluntary actions)
   - Proprioceptive feedback enabling awareness of body position and movement
   - Perception of action quality (fluent, awkward, comfortable, weird) through kinesthetic sense
7. **Perceptual Control**:
   - Axiom 47.32: Sense organ actions as movements to enable or prevent perception
   - Axioms 47.33-47.34: Tuning in/out through willing perceptions into/out of focus
   - Active control over attention and perception through voluntary mental actions
8. **Capability and Activity Levels**:
   - Axioms 47.35-47.39: Functional entities can be intact, impaired, or destroyed
   - Axioms 47.40-47.45: Mind states (active, impaired, inactive) based on focus and reasoning
   - Axiom 47.42: Active mind has thoughts in focus and reasoning processes occurring
   - Axiom 47.43: Impaired mind has thoughts in focus but no coherent reasoning
9. **Channel Theory**:
   - Axiom 47.46: Every person has sensation and control channels connecting mind and body
   - Axioms 47.47-47.50: Sensation channel states based on perception accuracy and reality correspondence
   - Axioms 47.51-47.54: Control channel states based on willing effectiveness in producing intended actions
   - Three-valued state system (active/impaired/inactive) for systematic capability assessment
10. **States of Consciousness**:

- Axioms 47.55-47.56: Consciousness states as specific combinations of component states
- Axiom 47.55: "Fine" state with intact body and active mind/channels
- Axiom 47.56: Unconscious state with intact body but inactive mind/channels
- Systematic framework for 16 different mind-body states (Table 47.1)

47.2 Technical Sophistication:

- **Reified Mental Actions**: Extensive use of primed predicates (will', perceive', concentrate', move') for eventuality-based mental and physical processes
- **Channel Architecture**: Novel formalization of mind-body interaction through bidirectional channels with systematic state classification
- **Defeasible Psychology**: Uses (etc) predicate for non-monotonic reasoning about perception, action, and mental causation
- **Functional Decomposition**: Bodies, minds, and channels as functional composite entities with intact/impaired/destroyed states
- **Scale Integration**: Integration with skill scales and capability measurements from other chapters
- **Causal Precision**: Careful distinction between direct causation (dcause) and general causation for mental processes

47.3 Complexity Distribution:

- Simple: 25 axioms (basic definitions, type constraints, classification schemes)
- Moderate: 26 axioms (mental state definitions, action taxonomies, channel characterizations)
- Complex: 5 axioms (coordinated actions, logical reasoning, sensation channel activity, control channel inactivity, attend to sense)

47.4 Conceptual Importance:

This chapter provides essential infrastructure for: - **Philosophy of Mind**: Formal treatment of mind-body dualism with computational precision - **Cognitive Psychology**: Models of awareness, attention, perception, and voluntary action - **Artificial Intelligence**: Framework for embodied cognition and mind-body interaction in agents - **Natural Language Understanding**: Lexical semantics for mind-body vocabulary (dexterous, clumsy, aware, numb, paralyzed) - **Human-Computer Interaction**: Models of perceptual control and action feedback - **Medical/Clinical Applications**: Systematic classification of consciousness and capability states

47.5 Cross-Chapter Connections:

- **Chapter 19 (Persons)**: Fundamental person, body, mind structure and perceive predicate

- **Chapter 21 (Knowledge)**: Focus of attention, memory, inm predicate, and perception-belief connection
- **Chapter 23 (Memory)**: Memory structure and focus/awareness distinction
- **Chapter 24 (Envisioning)**: thinkOf predicate and concentration mechanisms
- **Chapter 28 (Goals)**: Functional composite entities, intact/impaired/destroyed states
- **Chapter 30 (Pleasure and Pain)**: Bodily sensations and subjective experience
- **Chapter 41 (Planning)**: Plan execution framework for voluntary actions
- **Chapter 48 (Observation)**: skilled predicate and action quality assessment

47.6 Applications Mentioned:

- **Typing at terminal**: Focused on text but peripherally aware of overhead light
- **Reflex examples**: Knee-jerk response, blinking from sudden light
- **Voluntary control**: Directing gaze, cupping ears, flaring nostrils, holding nose
- **Attention control**: Tuning out distractions, focusing on distant conversations, ignoring pain
- **Coordination examples**: Multiple body parts working together with temporal coordination
- **Consciousness states**: Fine, drunk, paralyzed, unconscious, sleeping, brain dead

47.7 Notable Design Decisions:

- **Cartesian Dualism**: Adopts traditional mind-body dualism rather than contemporary functionalist or materialist views
- **Channel Mediation**: Mind-body interaction occurs through sensation and control channels rather than direct connection
- **Three-Valued States**: Systematic use of active/impaired/inactive for all mental and channel components
- **Reified Willing**: Will as eventuality that can directly cause bodily actions and mental states
- **Defeasible Perception**: Perception defeasibly leads to belief and focus, allowing for exceptions
- **Kinesthetic Integration**: Systematic treatment of proprioceptive feedback from bodily actions
- **State Systematicity**: 16 consciousness states from 4 components $\times$ 3 states, enabling precise distinctions

47.8 Theoretical Significance:

Chapter 47 represents the most comprehensive formal treatment of mind-body interaction in AI and cognitive science literature. By adopting commonsense dualism, it provides computational precision to folk psychological concepts while maintaining intuitive accessibility.

The channel theory innovation allows systematic modeling of mind-body interaction disorders: sensation channel problems (hallucinations, numbness) versus control channel problems (paralysis, clumsiness). This framework supports both normal and pathological states within a unified formalism.

The three-tier awareness architecture (memory-awareness-focus) provides computational structure for attention and consciousness while remaining grounded in commonsense psychology. The reified willing framework enables causal modeling of mental causation without falling into epiphenomenalism.

The systematic state classification enables precise lexical semantics for mind-body vocabulary, supporting natural language understanding systems that need to distinguish between "paralyzed" (intact body, inactive control channel) and "unconscious" (intact body, inactive mind and channels).

47.9 Philosophical Impact:

The formalization demonstrates that commonsense dualism can be given computational precision without philosophical incoherence. By treating willing as causally efficacious eventuality, it addresses traditional problems of mental causation in dualist frameworks.

The systematic treatment of consciousness states provides formal foundations for subjective experience categories used in clinical, legal, and everyday contexts. The framework supports reasoning about capability, responsibility, and mental health conditions.

The work shows how formal logic can capture the rich structure of folk psychology without reducing it to purely physical or computational terms, maintaining the explanatory adequacy of commonsense mental concepts while enabling systematic inference and reasoning.

This represents a significant contribution to both AI/cognitive science and philosophy of mind, demonstrating how formal methods can illuminate rather than eliminate commonsense psychological categories.

Chapter 48

Chapter 48: Observation of Plan Executions

- **30 axioms total** covering observation of plan executions, instructions, performances, skill assessment, and evaluation of agent behaviors
- **5 main sections**: Observing Plan Executions of Other Agents, Instructions, Performances and Their Specification, Skill, Evaluation
- **All psychology** - focuses on social aspects of plan execution, instruction following, skill assessment, and performance evaluation in multi-agent contexts

48.1 Key Features Identified:

1. **Observation Theory Framework**:
   - Axioms 48.1-48.7: Systematic distinction between observable vs. observed plan executions
   - Axiom 48.1: Observable execution requires at least one subgoal execution perceivable under constraints
   - Axiom 48.2: Unobservable execution has no perceivable subgoal executions under constraints
   - Axioms 48.3-48.4: Observed vs. unobserved executions based on actual perception rather than possibility
   - Axioms 48.6-48.7: Logical relationships between observability and observation (observed implies observable, unobservable implies unobserved)
2. **Communication and Instruction Theory**:
   - Axiom 48.8: Documents as meaning-bearing objects linking symbols to concepts for social groups
   - Axiom 48.9: Instructions as documents whose content is a plan for achieving goals
   - Axiom 48.10: Explicit plans as those with documentary specifications known to agents

- Integration with communication theory through (mean x y s) predicate relating symbols to concepts for social groups

3. **Performance Framework**:
   - Axiom 48.11: Performances as executions of explicit plans before audiences who perceive the execution
   - Axiom 48.12: Performance specifications as instructional documents for performances
   - Axioms 48.13-48.17: Complete lifecycle of performance specifications from candidate generation through validation/invalidation
   - Distinction between candidate specifications (believed possibly correct) and validated specifications (confirmed correct)
4. **Skill Assessment Theory**:
   - Axiom 48.18: Right way as executing plans according to their documentary specifications
   - Axiom 48.19: Attempted executions as executing plan p1 while intending to execute plan p for same goal
   - Axioms 48.20-48.21: Two conditions for "more skilled" based on subgoal matching (fewer extraneous, more correct subgoals)
   - Axiom 48.22: Skill scales defined by attempted executions ordered by moreSkilled relation
   - Axiom 48.23: Skilled executions as those in high region of skill scale
5. **Skill Level Systematization**:
   - Axiom 48.24: Skill levels as equivalence classes where equally skilled executions cannot be ordered
   - Axiom 48.25: Skill levels as "at" relations positioning executions on skill scales
   - Axiom 48.26: Agent independence - skill levels depend only on execution characteristics, not performer identity
   - Enables comparative skill assessment across different agents performing same tasks
6. **Evaluation Framework**:
   - Axiom 48.27: Evaluation as perceiving a performance and forming belief about its skill level
   - Axiom 48.28: Evaluation results as skill level judgments produced by evaluation events
   - Axiom 48.29: Witnessing attempted executions defeasibly causes evaluation (with etc condition)
   - Axiom 48.30: Evaluation criteria as properties of executions causally involved in evaluation results

48.2 Technical Sophistication:

- **Reified Mental Processes**: Extensive use of primed predicates (executePlan', perceive', evaluate', attemptExecute') for eventuality-based actions and mental states
- **Social Communication Theory**: Integration with communication framework through mean predicate relating symbols to concepts for social groups

- **Scale Theory Integration**: Skill scales with partial ordering (moreSkilled) and scale regions (Hi) for systematic skill assessment
- **Set-Theoretic Skill Analysis**: Sophisticated use of set operations (setdiff, intersection, subset, properSubset) for comparing subgoal achievement
- **Defeasible Social Psychology**: Uses (etc) for non-monotonic reasoning about evaluation following from observation
- **Agent-Independent Assessment**: Formal treatment ensuring skill judgments are objective rather than performer-relative

48.3 Complexity Distribution:

- Simple: 9 axioms (basic definitions, logical relationships, constraint specifications)
- Moderate: 15 axioms (multi-step processes, performance management, evaluation frameworks)
- Complex: 6 axioms (skill comparison with set operations, skill scale definitions, evaluation causation)

48.4 Conceptual Importance:

This chapter provides essential infrastructure for: - **Social Psychology**: Models of skill assessment, performance evaluation, and instruction following in social contexts - **Educational Technology**: Formal frameworks for skill assessment, performance feedback, and instructional design - **Human-Computer Interaction**: Models of system instruction, user performance evaluation, and skill-based adaptation - **Artificial Intelligence**: Multi-agent coordination through instruction giving, performance monitoring, and skill-based task allocation - **Organizational Psychology**: Formal models of training, performance evaluation, and skill development - **Natural Language Understanding**: Lexical semantics for skill and performance vocabulary (skilled, clumsy, expert, novice)

48.5 Cross-Chapter Connections:

- **Chapter 41 (Planning)**: Fundamental plan execution framework (executePlan', execute', subgoal relationships)
- **Chapter 21 (Knowledge)**: Belief formation (believe') and perception-to-belief processes
- **Chapter 19 (Persons)**: Basic perceive predicate and agent-action relationships
- **Chapter 12 (Scales)**: Scale theory infrastructure (scaleDefinedBy, Hi, inScale, at relations)
- **Chapter 15 (Causality)**: Causal relationships (cause, causallyInvolved) in evaluation and instruction
- **Chapter 47 (Mind-Body)**: Performance perception and skilled action concepts
- **Chapter 6 (Sets)**: Set operations for subgoal analysis and skill comparison

48.6 Applications Mentioned:

- **Microsociology**: Interactions among agents through observation of purposeful actions
- **Instruction following**: Document-based specification of plans and their execution
- **Performance evaluation**: Skill-based assessment of plan execution quality
- **Training and education**: Candidate performance specification generation, modification, validation
- **Expert systems**: Skill level assessment and performance criteria identification
- **Quality control**: Evaluation of task performance against standards

48.7 Notable Design Decisions:

- **Observable vs. Observed Distinction**: Careful separation of possibility (constraints) from actuality in perception
- **Document-Mediated Communication**: Instructions as meaning-bearing objects rather than direct plan transfer
- **Candidate Specification Lifecycle**: Complete process from generation through validation/invalidation
- **Set-Theoretic Skill Definition**: Formal skill comparison based on subgoal set analysis rather than qualitative judgments
- **Agent-Independent Skills**: Skill levels as properties of executions rather than performer characteristics
- **Defeasible Evaluation**: Non-monotonic reasoning about evaluation following from observation
- **Scale-Based Skills**: Integration with general scale theory for systematic skill level representation

48.8 Theoretical Significance:

Chapter 48 provides the most comprehensive formal treatment of skill assessment and performance evaluation in AI and cognitive science. By grounding skill in subgoal achievement rather than subjective judgments, it offers objective criteria for performance assessment while maintaining psychological realism.

The observation theory framework enables systematic reasoning about what aspects of others' actions can be perceived and how this relates to skill evaluation. The distinction between observable and observed executions provides foundations for reasoning about privacy, surveillance, and information sharing in multi-agent contexts.

The instruction and performance framework offers formal foundations for educational technology, training systems, and human-computer interaction. The candidate specification lifecycle models

how people develop, test, and validate their understanding of how tasks should be performed.

48.9 Philosophical Impact:

The chapter addresses fundamental questions about objective skill assessment by grounding skill in execution-plan matching rather than subjective evaluation. The agent-independence principle ensures that skill assessments reflect task performance quality rather than evaluator bias or performer identity.

The communication theory integration shows how symbolic instructions mediate between plans and their execution, providing foundations for understanding how knowledge is transmitted through documents, manuals, and educational materials.

The evaluation framework provides formal foundations for understanding how observation leads to judgment, with evaluation criteria as causally relevant properties of performances. This supports reasoning about fair assessment, evaluation bias, and performance feedback.

48.10 Social Psychology Contributions:

The work provides formal foundations for key concepts in social psychology including: - **Skill attribution and assessment** in professional and educational contexts - **Performance evaluation** and feedback in organizational settings
- **Instruction giving and following** in hierarchical relationships - **Expert-novice distinctions** based on formal skill scale positioning - **Social learning** through observation and evaluation of others' performance

The framework supports reasoning about social phenomena like performance anxiety (being observed), skill recognition, mentoring relationships, and professional development.

48.11 Educational Technology Applications:

The formal framework enables: - **Automated skill assessment** based on execution-plan matching - **Personalized instruction** adapted to skill level positioning - **Performance feedback systems** with objective skill criteria - **Learning analytics** tracking skill development over time - **Collaborative learning** environments with peer evaluation capabilities

This represents a significant advancement in providing computational foundations for educational technology that go beyond simple correctness checking to nuanced skill assessment and development support.

Chapter 49

Chapter 49: Emotions

- **120 axioms total** covering emotional intensity, basic emotions, cognitively elaborated emotions, liking/disliking, and emotional processes
- **10 main sections**: Intensity & Arousal, Happiness & Sadness, Shades of Happiness, Raw Emotions, Cognitively Elaborated Emotions, Hope & Fear, Reactions to Goals, Achievements & Failures, Envy & Jealousy, Liking & Disliking, Emotional States & Tendencies, Appraisal & Coping
- **Pure psychology** - comprehensive formalization of commonsense emotion theory bridging cognitive science and AI

49.1 Key Features Identified:

49.1.1 1. Emotional Intensity Framework:

- Axioms 49.1-49.10: Intensity based on goal importance, event size, and response magnitude
- Integration with scale theory using Hi regions and moreIntense relations
- Arousal as consequence of intense happiness, anger, and fear
- Arousal causes focus on current world understanding

49.1.2 2. Basic Emotions Theory:

- **Happiness & Sadness** (49.11-49.22): Goal-based emotions affecting action and belief change
- **Raw Emotions** (49.33-49.46): Fear, anger, disgust as threat responses
  - Fear: Cannot eliminate threat → avoid threat
  - Anger: Can eliminate threat → eliminate threat
  - Disgust: Interior threat → eject threat
- Sophisticated threat taxonomy based on location (interior/exterior) and agent capability

49.1.3 3. Happiness Varieties and Nuances:

- Axioms 49.23-49.32: Joyful, vindicated, pleased, glad, cheerful, jubilant, elated, euphoric
- Systematic treatment of semantic distinctions in happiness terms
- Social vs. individual manifestations (cheerful → social interaction)
- Intensity gradations (elated → euphoric as Hi region of Hi region)

49.1.4 4. Cognitively Elaborated Emotions:

- **In-group/Out-group Framework** (49.47-49.57): Shared vs. competitive goals determining emotional responses
  - In-group success → happiness for them, failure → sorrow for them
  - Out-group success → resentment, failure → gloating
- **Anticipation-Based Emotions** (49.58-49.73): Hope, fear-2, satisfaction, disappointment, relief
- **Achievement Emotions** (49.74-49.92): Pride, gratification, appreciation, gratitude, self-reproach, remorse, embarrassment, reproach, anger-2

49.1.5 5. Envy and Jealousy Theory:

- Axioms 49.93-49.98: Formal distinction based on goal exclusivity
- Jealousy: Mutually exclusive competitive goals (only one can win)
- Envy: Non-exclusive goals where you fail and they succeed
- Different emotional consequences: jealousy → anger, envy → sadness

49.1.6 6. Liking and Disliking Framework:

- Axioms 49.99-49.106: Dispositional emotional states based on happiness/unhappiness causation
- Bi-directional causation: liking causes happiness, happiness causes liking
- Extension from eventualities to entities via arg* relations
- Love/hate as intense versions of liking/disliking

49.1.7 7. Meta-Emotional Processes:

- **Emotional States** (49.107-49.116): Classification system, change processes, tendencies
- **Appraisal Theory** (49.117-49.118): Belief change causing emotional states

- **Coping Theory** (49.119-49.120): Strategies for managing unhappiness

49.2 Technical Sophistication:

- **Extensive Defeasibility**: 66 axioms use (etc) - highest defeasible content showing emotion's inherent non-monotonic nature

- **Reified Emotional States**: Systematic use of primed predicates (happy', sad', angry', etc.) enabling temporal and causal reasoning
- **Scale Integration**: Sophisticated use of intensity scales, Hi regions, and comparative relations
- **Social Cognition**: Complex in-group/out-group distinctions affecting emotional responses
- **Threat Taxonomy**: Systematic classification of threats by location and agent response capability
- **Anticipation Framework**: Future-oriented emotions based on graded belief and envisioning

49.3 Complexity Distribution:

- Simple: 42 axioms (basic emotional implications, type constraints, definitional equivalences)
- Moderate: 60 axioms (standard emotional causation, social emotional responses, defeasible rules)
- Complex: 18 axioms (sophisticated cognitive processes, multi-level threat analysis, recursive emotional structures)

49.4 Conceptual Importance:

This chapter provides crucial infrastructure for: - **Affective Computing**: Computational models of emotion for human-computer interaction - **Social Robotics**: Understanding emotional responses in social contexts - **Cognitive Architectures**: Integration of emotion with reasoning, planning, and learning - **Computational Psychology**: Formal models of human emotional processes - **AI Safety**: Understanding human emotional responses to AI systems and actions

49.5 Cross-Chapter Connections:

- **Chapter 5 (Eventualities)**: Fundamental eventuality framework for emotional states
- **Chapter 12 (Scales)**: Intensity scales and comparative emotional relations
- **Chapter 15 (Causality)**: Extensive causal relations in emotional processes
- **Chapter 20 (Modality)**: Graded belief in anticipation-based emotions
- **Chapter 21 (Knowledge)**: Belief change in emotional appraisal
- **Chapter 28 (Goals)**: Goal framework underlying most emotional responses

49.6 Applications Mentioned:

- **Threat Assessment**: Bear encounter (fear vs anger based on elimination capability)
- **Social Emotions**: Happiness/sadness for in-group vs out-group members
- **Achievement Emotions**: Pride in children's accomplishments, gratitude for help received
- **Anticipation**: Hope for lottery winnings, fear of being late
- **Aesthetic Emotions**: Satisfaction from goal achievement, disappointment from failure

49.7 Notable Design Decisions:

- **Appraisal Theory Integration**: Emotions caused by belief changes about environment
- **Threat-Response Framework**: Systematic classification of fear, anger, disgust by threat type
- **Social Categorization**: In-group/out-group distinctions determining emotional responses
- **Defeasible Causation**: Extensive use of (etc) recognizing emotion's context-dependent nature
- **Intensity as Scale Position**: Integration with mathematical scale theory for comparative emotions
- **Cognitive Elaboration**: Complex emotions as variations on basic five (happiness, sadness, anger, fear, disgust)

49.8 Theoretical Significance:

Chapter 49 represents the most comprehensive formalization of commonsense emotion theory in the AI literature. The systematic treatment of emotional intensity, the sophisticated threat-response taxonomy for raw emotions, and the detailed analysis of social emotions provide both philosophical rigor and computational tractability.

The extensive use of defeasible reasoning (66 axioms with etc) reflects the inherently context-dependent nature of emotional responses, while the integration with scale theory enables quantitative reasoning about emotional intensity and comparison.

The in-group/out-group framework provides a foundation for understanding social emotions crucial for multi-agent systems and human-AI interaction. The anticipation-based emotions integrate planning and temporal reasoning with affective states.

This chapter bridges multiple disciplines - cognitive psychology, affective computing, social cognition, and AI - providing formal tools for reasoning about one of the most important aspects of human experience. The 120 axioms establish comprehensive coverage from basic biological responses to sophisticated social emotions, enabling AI systems to understand and respond appropriately to human emotional states.

The theoretical framework supports both recognition of human emotions and generation of appropriate emotional responses in artificial agents, making it essential for socially intelligent AI systems operating in human environments.