

A Model of Poetic Comprehension

Kenneth Haase

MIT Media Laboratory
20 Ames Street
Cambridge, Massachusetts 02139
haase@media.mit.edu

Abstract

This article introduces an account of aesthetic comprehension and experience together with an implemented miniature which generates analogical interpretations from a semi-automatic parse of Wordsworth's "Lines Written in Early Spring". In our account, a poem serves as an *analogy teaching machine* by using formal structure to cue the formation of novel analogies. This account builds on an analogical model of comprehension previously applied to large corpora of newspaper summaries. In the miniature, an automatic grammatical and semantic analysis of the text is augmented with information about rhyme and rhythm. These formal cues allow the system to determine analogies which it would not otherwise consider. The article describes the comprehension framework, the annotated piece, and the matcher's performance on the piece. It closes with a discussion of possible objections to aspects of the thesis or experiment and suggested directions for future work.

Introduction

This article introduces an account of aesthetic comprehension and describes an implemented miniature inspired by the account. The account begins with a general model of comprehension where routine and systematic analogies among descriptions takes the place of translation into canonical form. In this framework, semantic correspondence is based on dynamic analogizing rather than structural or nominal alignment of canonical descriptions. In other work, this model has been applied to indexing and analyzing a large (~ 10 million word) corpus of short news summaries (Haase 1995).

In this article, we discuss the application of the same mechanisms to a transcription of "Lines Written in Early Spring" by William Wordsworth (1770-1850). Our thesis is that aesthetic experience involves the identification of new analogies and similarities and that the formal structure of a piece provides the cues which enable such analogies to be considered. Given annotations describing rhythmic and rhyming structure, our analogical matcher constructs mappings consistent (in many cases) with the metaphor

Wordsworth is invoking. In our (partially implemented) comprehension framework, these associations enable the future identification of similar analogies in the absence of formal cues. To follow on Minsky's analysis of musical understanding (Minsky 1981), we propose that a poem functions as an *analogy teaching machine*.

Our miniature illustrates a characterization of aesthetic experience as involving certain radical reorganizations of memory based on the consideration of new analogies. We propose that aesthetic experiences change the way our minds match and index subsequent experiences. This account assumes that our daily experience is organized around a dynamic memory (Schank 1982) and that aesthetic experiences are those which change the indexing and matching strategies of this memory. Though this may seem an oversimplification if we consider the small dynamic memories our programs have had to date, it seems less objectionable if we consider the kinds of rich dynamic memories which humans really have, accumulating years of lived experience.

In the rest of this article, we introduce the analogical comprehension model, describe the representations of prose and poetry it operates over and discuss its performance on Wordsworth's *Lines* and its sensitivity to different formal cues (rhyme, rhythm, etc.). We then discuss how the system may operate differently in the future based on its partial comprehension of *Lines* and lead into a discussion of the general model of aesthetic experience which we are proposing.

Analogical Comprehension

This section sketches the model of analogical comprehension applied to the interpretation of *Lines*. Analogical comprehension replaces the reduction to and comparison of canonical forms with the determination of dynamic analogies between individual non-canonical descriptions.

Descriptions in this framework consist of nodes in a semantic network connected to one another by two kinds of relations: **micro-relations** capture the *structure* of individual descriptions; **associations** capture the *significance* of individual nodes by connection either to

individuals in other descriptions or to nodes in a shared ontology. In the description of prose text, for instance, *sentences* are descriptions, *phrases* are nodes, *grammatical structure* is represented by micro-relations, and *possible word meaning* is encoded as association relations.

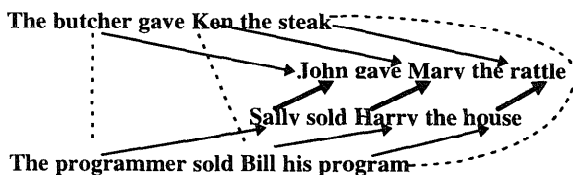
Descriptions are matched at two levels:

- cognate matches are based solely on associations
- structural matches extend cognate matches based on micro-relational structure

Cognate matches are pairings having some unique common association with respect to their contexts. For example, in the descriptions of the sentences “Chris embraced Terry” and “Pat hugged Robin” where each word is a node, the nodes representing “embraced” and “hugged” would be cognates because they have a common association (e.g. “embrace, hug, bosom, or squeeze”) in WordNet) shared by none of the other nodes. The other possible pairings, however, would not be cognates because any associations they have in common (e.g. “person”) are common to all of them.

Structural matches extend cognate matching based on structural systematicity (Gentner 1983) of micro-relations. In the example above, subject and object relations of identified cognates (“embraced” ⇒ “hugged”) generate the mappings (“Chris” ⇒ “Pat”) and (“Terry” ⇒ “Robin”). Such extensions remain constrained associational information, requiring some common association between the linked nodes. When micro-relational structure is ambiguous (e.g. the subject of a verb phrase is uncertain), *unique* common associations are required with respect to the candidates.

The overall comprehension framework includes a (still experimental) indexing facility which associates each new description with previously encountered descriptions of similar associational and micro-relational structure. In this framework, cognate matches through these structural prototypes will reflect structural roles as well as term similarity. Association of structural roles with each other can then reflect semantic similarity of structural roles, as in:



where particular semantic associations (the heavy lines) combined with automatic structural associations (the light lines) yield semantically significant cognate relations (the dashed lines and arc). From this example, we can see that the formation of prior semantic associations (the heavy lines) is the basis for subsequent comprehension. Our model of aesthetic comprehension is an account of the role aesthetic experience plays in the formation of some of these prior analogies.

Comparison to other analogical matchers

Unlike the base level of matching in SME (Falkenhainer, Forbus, and Gentner 1989), the cognate relation is contextually sensitive and can be changed without modifying or extending the matcher itself. Unlike ACME (Holyoak and Thagard 1989) or CopyCat (Mitchell 1993), where matching is also contextually sensitive, the cognate relation is also defeasible: a match depends on the existence of unique common associations which can be added or removed rather than on a combination of weights and activation levels.

Representing Poetry

The representation of poetic text extends the representation of prose text with associations based on rhyme and meter. We retain the grammatical micro-relations and add associations for each node (phrase) representing:

- possible meaning, based on WordNet (Miller 1990)
- final phoneme (representing rhymes)
- metrical position in the line (1st beat, 2nd beat, etc.)

The meaning representation starts with a node based on surface form and part of speech for each word in the phrase. This is in turn associated with all of the WordNet senses for that combination and the WordNet senses are then associated with their hypernyms (generalizations). For instance, the phrase “in a grove” is associated with nodes **in.prep**, **the.det**, and **grove.noun**; **grove.noun** is then associated with WordNet synsets for “grove” and “grove, woodlet, or orchard”. These are then associated with their respective hypernyms: “forest, wood, or woods” for “grove” and “garden” for “grove woodlet or orchard” and so on for their hypernyms (generalizations)..

The representation explicit preserves ambiguity in both grammatical structure and word meaning. In prose understanding, this allows interpretation to be delayed until disambiguating contextual cues are available, much as in (Hirst 1987). In poetic interpretation, this provides the “play” which allows metaphorical interpretations to emerge.

Representing Lines

To our pleasant surprise, our parser did a passable job of analyzing the Wordsworth poem, which consists of six stanzas of two couplets each. The couplets were treated as sentences and passed to our parser. The parser produced a node for each phrase and these were then annotated with associations based on rhyme, rhythmical position, broad syntactic category (thing, action, etc.) and possible word meaning (representing ambiguously via WordNet). A node representing a phrase was counted as being on a beat if its head (e.g. the noun or verb) fell on the beat, leaving some beats unaccounted for. Some analogies were precluded by the use of a phrase-level representation; e.g. a promising

match between “blended” and “pleasant” in the first stanza was not recognized because the phrases, rather than the words, were distinguished as individuals.

The Engine Matches

Given this representation, our program takes each represented stanza and attempts to match its two couplets.¹ This particularly suits *Lines* (though we didn’t realize this until examining the program’s analysis) because each stanza is divided into a ‘naturalistic’ and ‘contemplative’ couplet. The matcher determines analogies which sometimes (but not always) fits the mapping of mental and natural realms which Wordsworth is trying to invoke. In the listing to the right, we show each stanza of the poem annotated with the information available to the matcher. Alternating italic and roman text indicate phrase structure, subscripts indicate metrical position of a phrase’s head, and underline words indicate rhymes. Some rhymes are duplicated within a stanza, particularly “Man” in the second and last stanza. To the right of each stanza are the matches the system found with cognate matches (in italics) listed first and structural matches listed subsequently.

The described matches were based on all the sorts of associations described above, but in order to determine the role which different annotations played in the final match, we processed the text several times while selectively disabling different kinds of association (e.g. ignoring rhyme). We will refer to some of these variations in our discussion of the stanzas². Note that adding new associational information can either add or remove existing matches by introducing new connections or making existing connections ambiguous.

In the first stanza, cognate matching pairs “reclined” and “mind” (silly, but based on meter and rhyme) and “grove” and “mood” (more interesting and based on their common role as “settings”). Structural extension of this second match yields a pairing of “heard” and “bring” which in turn yields a match between “notes” and “thoughts”. This trio of matches seems consonant with Wordsworth’s intended juxtaposition of nature and thought in the poem.

In the second stanza, as in the first, meter and rhyme cue an initial analogy between “link” and “think”; the rhyme between “ran” and “Man” doesn’t result in an initially analogy because the second occurrence of “Man” keeps the rhyme from being unique. Both rhythm and WordNet suggest the link between “Nature” and “heart”. The connection between “did” and “grieved” doesn’t make a lot

¹The actual mappings and transcripts of the program can be found on the World Wide Web at

<http://mu.www.media.mit.edu/projects/poetry>

I heard ₁ a thousand blended notes ₄ While in a grove ₆ I sat ₇ reclined ₈ In that sweet mood ₂ when pleasant thoughts ₄ Bring sad thoughts to the mind ₇	Grove ⇒ mood reclined ⇒ mind heard ⇒ bring notes ⇒ thoughts
To her fair works ₂ did Nature ₃ link ₄ The human soul ₆ that through me rang ₇ And much ₁ it grieved ₂ my heart ₃ to think ₄ What Man ₅ has made ₆ of Man ₇	Nature ⇒ heart link ⇒ think did ⇒ grieved soul ⇒ Man ran ⇒ made
Through primrose tufts ₂ in that sweet bower ₄ The periwinkle ₆ trailed ₇ its wreaths ₈ And ‘tis my faith ₂ that every flower ₄ Enjoys ₁ the air ₁ it breathes ₁	
The birds ₁ around me hopped ₃ and played ₄ Their thoughts ₅ I cannot measure ₇ But the least motion ₂ which ₃ they made ₄ It seemed ₅ a thrill ₆ of pleasure ₇	birds ⇒ motion played ⇒ made thoughts ⇒ seemed
The budding twigs ₂ spread out their fan ₄ To catch ₅ the breezy air ₇ And I ₁ must think ₂ , do all I can ₄ That there ₅ was pleasure ₆ there ₇	twigs ⇒ think fan ⇒ can catch ⇒ there air ⇒ pleasure
If this belief ₂ from heaven ₃ be sent ₄ If such ₅ be Nature’s holy plan ₈ Have I ₁ not reason ₂ to lament ₄ What Man ₅ has made ₆ of Man ₇	belief ⇒ reason sent ⇒ lament such ⇒ Man ₅ plan ⇒ Man be ⇒ made

The Poem and Its Matches

of sense but is justified by a relatively obscure path through WordNet. The connection between “soul” and “Man” is striking and is based on the proximate link of the verbs “link” and “think”. It in turn generates the link between “ran” and “made” which has a vague sort of thematic consistency though not as striking as some of the other metaphors determined elsewhere.

The third stanza is singularly unproductive; the problem is that there are no cognate relations to start with because there are too many unique common associations. The link between “bower” and “flower” based on rhyme and rhythm competes with a link between “tufts” and “flower” based on WordNet while metrical and rhyme matches are also in conflict. As we take information away, certain mappings emerge, but none of them are very striking; most are the consequence of straightforward alignment of meter or rhyme. Part of the problem is also that this stanza mixes the descriptions of nature and thought; this may be intentional on Wordsworth’s part, but it’s not something which simple couplet to couplet matching can handle.

In the fourth stanza, despite a similar mingling of action and thought, the system does establish an interesting

correspondence between the birds in the first couplet and the motion he perceives in the second. All the connections here are cognates and the overall analogy does not have any satisfying systematicity.

In the fifth stanza, the links are all cognate relations based on rhythm or, in the case of “air” and “pleasure” on the WordNet synset for “activity or behavior,” which is quite a stretch (as in ‘air one’s views’ and ‘pleasure oneself’).

However, after a dry spell, the matcher does a better job on the final stanza, where WordNet, meter, and part of speech conspire to create the map between ‘belief’ and ‘reason’ and meter and rhyme establish the other initial mappings. The initial mappings seem to make a certain narrative sense; the active actions (sent and lament) coincide and the match between “belief” and “reason” is consistent with the metaphor Wordsworth has been using. The generated match between “plan” and “Man” might be reflect Wordsworth’s expression of sadness at some lost potential and the link between “be” and “made” suggests that this loss is somehow intentional.

In brief summary, the matcher produced relatively reasonable matches between the couplets of stanzas 1, 2, and 6; it did little of interest with 3, 4, and 5, possibly because they lacked significant inter-couplet structure.

Aesthetic Function

How do the matches determined in our miniature relate to our model of aesthetic comprehension? The answer lies in how new situations are processed in the analogical comprehension model. Briefly, when a new description is encountered, the system searches for previous descriptions with similar associations and micro-relations; it then attempts to analogize between the new description and these previous descriptions and these analogies become associations fixed in memory. Consider the analogies determined in the first stanza between ‘Nature link’ (1) and ‘Heart think’ (2); given these analogies fixed as associations, consider the case where two new descriptions arrive: ‘thunderstorms cause’ (3) and ‘leaders decide’ (4). These might be associated (through search and analogy) to (1) and (2) based on the similarity of causing and linking and thinking and deciding. However, through this connection and the precedent connection of (1) and (2), an association between the (3) and (4) might be determined based on the analogy and association established by exposure to the poem. The point here is not that such an association is always valid or sensible, but that the exposure to the match of (1) and (2) enables the consideration of a match between (3) and (4) even in the absence of strong rhythmic or rhyming cues. The analogies set up in stanza 1, in this case, enable future cases to be seen differently. This is the core of our account of aesthetic experience: it makes us see things differently by changing the structure of our memory. This work has shown how

artistic form can introduce analogies; the tantalizing hypothesis that this changes future comprehension is yet to be shown. To implement the example above, while possible, would be a contrivance: the real test must come in indexing and matching against a larger corpus.

Anticipated Objections

This section discusses the previous sections by considering objections which might be made to either the question itself, the approach taken, or the results and their significance.

Objections to the question

Two obvious objections to the very enterprise begun in this article come from two different camps. One is a humanist objection that to “reduce” aesthetic experience to some computational model will rob it of its power and make the world a much poorer place in which to live. This is certainly true in the sense that any critical analysis of an experience ‘takes us out of’ the experience and thus diminishes it in certain ways. On the other hand, such analysis can also enrich experience in other ways. “To explain” and “to explain away” are not necessarily the same thing. While it is vitally important to maintain respect for aesthetic experience, analysis need not be disrespectful. Indeed our thesis is that aesthetic experience is basically about changing the way the world is seen, according it a central role in human understanding.

A different objection may come from colleagues concerned that trying to understand aesthetic experience is “setting our sights too high”. Such concerns might be phrased thus: “get a handle on conventional reasoning and cognition and then start worrying about aesthetic experience; understanding literal meaning first and then start worrying about metaphor!” This concern is a valid one if aesthetic experience in fact builds upon a foundation of literal understanding and everyday problem solving. However, if the dependency goes the other way or if the two phenomena are co-dependent, understanding aesthetic experience is both comparable to and necessary for understanding everyday cognition. Indeed, given the model of aesthetic experience as memory reorganization, “deep learning” (of representations, for instance) may always be an aesthetic experience.

Objections to the approach

There can be little argument that the choice of poetic understanding and of *Lines* in particular was somewhat contrived. Our focus on poetry came from the existence of parsers and matchers built for handling prose and also from the accessibility of the medium to a wide audience of readers. *Lines* was not selected entirely at random: dozens of poems were looked at with an eye towards pieces with rich rhythmic and rhyming structure to provide cues for matching. There were roughly half a dozen candidates

originally selected and *Lines* was chosen more or less arbitrarily from this set. Because annotation of rhyme and meter had to be done by hand, a single poem was selected at this point.

The need for structure reflects another possible objection: human understanding of a poem plays on a rich articulated background of experience which any program necessarily (at this point) lacks. WordNet, for all its scope and detail, is a poor substitute for the lived life which humans bring to bear when reading poetry. Our system would be confounded by less structured verse, where the connections and associations rely on understanding text with respect to experience. Interpreting 'shall I compare thee to a summer's day, thou art more lovely and more temperate' requires knowing properties of summer and properties of lovers instead of mere associations of words. Such knowledge is as often experiential and poetry often evokes such common experience as well as formal structures to establish the analogies it teaches.

Objections to the results

Our results are of two sorts: the mappings determined between stanzas in *Lines* and the way in which these mappings influenced future understanding. I'll address some possible objections to each of these in order.

The problem with evaluating the matches determined by the system is in the character of poetic language: there is no 'objective' standard to hold the program's performance against. For instance, the sensibility of the mapping between 'Man' and 'plan' in the last stanza depends on an interpretation of the second occurrence of 'Man' as denoting 'What Man could have been'. Nonetheless, the goal is not to get the "correct" mapping because there isn't one; instead, the goal is to get a mapping which is plausible.

The second substantial problem is the argument about how exposure to *Lines* changed the analogies which the system would draw in the future. As a demonstration, the single case is unconvincing: in particular, it is not clear that the 'new perspective' generalizes (applies to many different descriptions) or doesn't over-generalize (apply to everything in sight). However, the case is mostly intended to be illustrative rather than demonstrative: the real test comes in a rich background with a larger corpora of understood 'real-life' texts. We cannot do this currently because our text database does not have the intra-textual mappings which would enable aesthetic experiences to transform them. We look forward to examining this question in the future.

Objections to the thesis

Our definition of aesthetic experience is that it changes the way in which our memories are indexed and matched. It is important that this definition be at the right level: it should not exclude some aesthetic experiences nor should it

include experiences which are not aesthetic. Because this is an inductive question, new cases may always break it, but in this section I address some obvious ones.

Non-aesthetic memory reorganization. Are our memories ever reorganized without being an aesthetic experience? First, let's distinguish reorganizing changes to memory from mere 'additive' changes to memory. For instance, the experience of being served by a new person at the cafe does not generally (by my definition) reorganize memory. The following definition may make this clear: a change to memory counts as a reorganization when two previously similar or dissimilar experiences change their relation (i.e. become dissimilar or similar) by virtue of the change. Experiencing the new server at the coffee shop does not normally count as aesthetic, but if the new server is a former congressman or department head, it might!

One related objection takes the form 'a mugger certainly changes the organization of my memory, but I wouldn't call it an artistic experience!' This criticism is somewhat allayed by introducing the distinction between aesthetic and artistic experience: artistic experience is an intentionally aesthetic experience. When you're frightened by a mugger, it may be an aesthetic experience but it is certainly not artistic. However, when you're frightened by a Spielberg film, the experience is both aesthetic and artistic. Both cases make the 'dark of night' look different by changing the memories which are triggered and the interpretations placed on shadows. But Spielberg is seeking that effect (and affect) while the mugger just wants your money (actually Spielberg might also want some of your money, but that's not the point).

Aesthetic experience without reorganization. Can aesthetic experiences exist which do not reorganize our memories? When I hear 'Blackbird' (Lennon & McCartney) playing in the background for the 59th time (in my life), it is probably not reorganizing my memory, but it might still be an aesthetic experience. However, I would argue otherwise.

It is important to distinguish properties of the piece and properties of the experience. Experiences are aesthetic while pieces are artistic. In particular, pieces are artistic by virtue of their intent and the aesthetic experiences they invoke. The first time I heard 'Blackbird,' it was an aesthetic (and artistic) experience. However, as it plays in the background as I type against a deadline, it is certainly pleasant but probably not 'aesthetic' --- I don't have the time or attention to experience it aesthetically.

Aesthetic Feeling. Am I misappropriating the word 'aesthetic' from its feeling-based roots by my definition? We say that experiences are 'aesthetically pleasing' --- can we talk about an experience being aesthetic without discussing a corresponding feeling? Well, "aesthetic" is not necessarily tied to "pleasing": a given 'good' and terrifying scene (for instance, the shower scene from Hitchcock's

'Psycho'), has a clear aesthetic dimension, but 'pleasurable' would certainly not be an applicable predicate for anyone I'd care to know. Some of us can also, I argue, draw a certain aesthetic sense from a clever proof or abstract painting without any associated emotion. Our use of the word aesthetic, given these cases, must in some way be distinguished from the emotions associated with it.

Of course, aesthetic experiences are often connected with emotion because emotion is a particularly potent device for invoking and changing the structure of our memories. (Gelernter 1994) proposes that emotion (through "affect linking") drives much of the 'low focus' indexing and analogizing that underlies art and creativity. Likewise, (Ortony 1988) suggests that emotions provide cues for learning. For this reason, the experiences which access and then change our memory structures often invoke pleasure, fear, or other emotions. In addition, the experience of memory reorganization has a certain "felt" character of its own; it is not entirely (or even primarily) an intellectual experience but it does change the world (as we see it) and such changes often bring their own emotional charge quite independently of the events which bring about the change.

Future Work

In the final response to possible objections, we wandered far from a relatively small program matching skeletal descriptions of parsed poems. Future work will strike out across this gap, seeking to expand the performance of the program, the background it works against, the range of forms to which it is exposed, and a serious assessment of how the structures created by aesthetic comprehension change comprehension and performance in everyday life and problem solving.

Two near-term steps are applying these mechanisms to more pieces and the automation of the rhyme and rhythm annotation of parses. If the second were accomplished, it would be interesting to do matching in larger corpora of poems, among the works of a particular poet or across poets within a school or style. It would be interesting if exposure to some works in a genre enabled analogies in other works. Finally, in order to assess the role which these mappings play in affecting "everyday comprehension," we hope to have a version of our 'news database' indexed together with our 'poetry database' in order to look at the deep or fanciful mappings generated by their combination.

Applying these techniques to other artistic forms is another area of future work. The native representation of ambiguity we use may suit the overlapping and ambiguous aspects of musical structure. (Ruttenberg 1994) proposes to use methods like these to analyze musical pieces. Visual and plastic arts are more complicated, because of the increased difficulty of parsing.

Another interesting area of research would be to consider critical theory and practice in light of these technologies

and results. (Holyoak 1984) suggests that critical theory might be informed by efforts in analogical reasoning; this work can be considered a step towards such a connection.

Another open and interesting question is whether we can use this characterization of aesthetic experience to help us design aesthetic experiences which work in new interactive media. Most new media start by copying existing media until they discover the new effects which they alone can produce. If we better understand how new modes of experience connect to the organization of memory we may be better able to use those modes in the construction of rich aesthetic experiences.

Acknowledgments

This work builds on research supported by the News in the Future consortium at the MIT Media Laboratory. I would also like to thank Jay Keyser, Roz Picard, Whitman Richards, and Janet Cahn for lively discussion and insightful comments on both this work and its presentation.

References

- Falkenhainer, B., Forbus, K., and Gentner, D., "The Structure-Mapping Engine: Algorithms and Examples", *Artificial Intelligence* (41), 1989
- Gelernter, D., *The Muse in the Machine*, The Free Press, Macmillan, 1994
- Gentner, D., The Structure-Mapping Engine, *Cognitive Science*, 7 (2), 1983
- Haase, K., Analogy in the Large, Proceedings of IJCAI-96, MIT Press 1995.
- Hirst, G., *Semantic Interpretation and the Resolution of Ambiguity*, Cambridge University Press, 1987.
- Holyoak, K. and Thagard, P. "A computational model of analogical problem solving," In *Similarity and Analogical Reasoning* (edited by S. Vosniadou and A. Ortony), Cambridge University Press 1989
- Holyoak, K., An analogical framework for literary interpretation, *Poetics* 11:105-126, 1984.
- Miller, G. A. 1990, "WordNet: An On-line Lexical Database," *International Journal of Lexicography*, 3(4).
- Minsky, M., "Music, Mind, and Meaning," in *Computer Music Journal*, Volume 5, No. 3, Fall 1981
- Mitchell, M., *Analogy Making as Perception*, MIT Press 1993
- Ortony, A., Clore, G., and Collins, A., *The Cognitive Structure of Emotions*, Cambridge University Press 1988
- Ruttenberg, A. Memory Representation and Deployment in Musical Thinking, Unpublished PhD Thesis proposal.
- Schank, R., *Dynamic Memory*, Cambridge University Press, 1982.