Critical Creative Experiences in Learning and Poetry

Jerry R. Hobbs
Artificial Intelligence Center
SRI International

I.

This paper is very speculative. I attempt to elucidate somewhat two critical experiences in learning and two critical effects in the reading of poetry, in a way that brings out their similarity. The thrust of my argument is that two of the fundamental experiences in the reading of poetry instantiate on a small scale two central features of theory construction in individual learning, and, indeed, in the progress of scientific understanding of the world.

These critical moments are moments of creativity, in that something new and useful has been created in the structure of the knowledge we use to understand the world.

The first critical moment in learning occurs when, after we accumulate more and more facts about a particular domain, each placing the system of knowledge as a whole under more stress, suddenly things settle. We achieve a new, more economic way of viewing the domain. We see the pattern. The theory has become compact, in a moment of clarity. What before were isolated facts now become deductive consequences of suddenly understood underlying principles.

The corresponding critical moment in poetry occurs when after attempting to see the underlying coherence in a densely cohesive text, attempting to place into a single framework ideas and images that seem to be in conflict, we suddenly see the underlying coherence by combining normally distant concepts in our system of beliefs, or by being forced into a reanalysis of phenomena hitherto understood only simply, now to be understood in more complex ways.

The other experience in learning and poetry is described below, after some background has been built up.

These critical moments are examined in terms of the impact they have on our knowledge base or system of beliefs. Thus, some preliminary remarks are needed about the account of knowledge used in the analysis.

I am working from a position that accepts the common view that has emerged recently in artificial intelligence and cognitive science that people come to know or comprehend the world by constructing theories that, in the long run, best explain the observable evidence presented by perception. There are two processes here—interpretation and learning. First interpretation: My view is that we interpret our environment by coming up with the best explanation for the observables that the environment presents us with—C. S. Peirce's notion of abduction (Peirce, 1955). For example, if I see the branches of a tree outside my office window moving back and forth, I come up with the explanation that the wind is blowing. This might not be true. There may be a gardener below window level shaking the tree, but the most likely explanation, the best explanation, is that the wind is blowing.

This account of interpretation depends on our having a knowledge base or belief system consisting of a number of general rules or explanatory principles, such as that the wind causes the branches of trees to move, or that the wind tends to cause anything to move and the roots of a tree keep the base of the tree in one place. But the existence of a knowledge base means that the knowledge has to be acquired. We have to learn new facts and explanatory principles.

The challenge in coming to an understanding of learning is twofold: We must come up with a model in which plausibly incremental additions and modifications can be made to our knowledge base, to accommodate the ordinary learning that we engage in every waking moment of our lives. And we must come up with an account of those critical moments of clarity and insight in learning when our knowledge becomes, in a sense, more compact. Before I attempt to address these two problems, however, I need to go one step deeper into the mire of the particular framework I am assuming.

I adopt a logicist position, as it is known in AI, a highly productive framework that nevertheless is rather out of favor these days. The AI logicist position is that thought can be modelled as the manipulation of data structures and that these data structures correspond to logical formulae. In principle, although by no means yet in practice, a machine that works in that fashion can exhibit the same input-output behavior as an intelligent being. If we were to succeed, we could say that, for all we know, this is how the intelligent being works. This enterprise, in a sense, is the exploration of the space of possible ways in which an intelligent being could work.

The more abstract the level at which the enterprise proceeds, the more possibilities any given account will cover, and therefore the more likely it will be to be a version of the correct account. This is the justification of the use of logic. Brains represent and manipulate information, and the best tool we have for representing information in a precise, formal fashion is logic. This notion, that "logic is (or might as well be) the language of thought", should not be as controversial as it is. It amounts only to the assumption that the way brains store and use information has certain unsurprising properties:

- Conjunction: Two different thoughts can have an additive effect. It is possible to think two things at once.
- Exclusive disjunction: We recognize obvious contradictions. We know that dogs aren't cats.
- Modus Ponens: Given one thought and a general principle associating that thought to another, we can consider the second thought. That is, concepts are associated.
- Predicate-argument relations: We can know when two things are related, and we can label that relationship. We can think about entities having properties. The basic unit of information is a **proposition** consisting of a **predicate** applied to one or more **arguments**.
- Universal instantiation: We know general principles and we can recognize distinct instantiations of these general principles.

These five aspects of information, plus one more, constitute the essentials of the logic that I find it productive to work in.

The final feature is defeasibility. Whereas in standard logic, the correct application of a rule of inference to a **true** premise **always** results in a **true** conclusion, in the (nonmonotonic) logic that is needed in AI, the correct application of a rule of inference to a **good** premise **usually** results in a **good** conclusion. Every AI researcher has his or her own favorite mechanism for defeasible reasoning, and I'm no exception. For me it is "cost-based abduction", which is deduction plus making consistent assumptions at a cost, where one tries to minimize the cost of proofs by minimizing the assumptions (Hobbs et al., 1993)

One of the principal ways of minimizing the cost of proofs is to have one assumption explain more than one observable. For example, if I hear a loud noise and the lights go out, I could assume that something fell and that caused the loud noise, and that there was a power failure at the same time and that caused the lights to go out. Or I could assume that something fell against the power supply causing both the loud noise and the lights to go out. The latter is a more economical explanation, and is preferred. It gives us more bang for the buck, where the observables are bang and assumptions are the buck.

In this framework, knowledge and beliefs are represented as logical axioms, which have the form of a conjunction of propositions implying a conjunction of propositions, where each propostion is a predicate applied to one or more arguments:

$$(\forall x, z) p_1(x) \wedge p_2(x, z) \supset (\exists y) q_1(x, y) \wedge q_2(y)$$

An agent's knowledge consists of a very large collection of such axioms.

In logic, the term for a set of axioms explicating some domain is "theory", and so when I talk about our **theory** of a particular domain, I will mean it in the loose sense of what we know about that domain, but also in the technical sense of a set of axioms that explicate that domain.

This framework is specified in precise enough detail that we can be precise and detailed about what learning might consist of. Learning is the accumulation and modification of axioms, in a way that causes the agent to interact with the environment in a more appropriate or successful manner. These modifications normally happen incrementally, so it is necessary to specify what sorts of incremental changes can occur, and since I have said that knowledge is represented as collections of axioms expressing inferential relations among conjunctions of propositions, I can use this characterization to drive the analysis of how axioms can change. I think that we can identify at least seven types of incremental change.

1. We learn new particular facts. We meet John, whom we have heard of, in person for the first time and see that he is tall, and the proposition $tall(John)$ becomes part of our knowledge base.

2. We refine our knowledge by adding a proposition to the antecedent of a general rule. A child who thought that anything one can sit on is a chair

   $$seat(x) \supset chair(x)$$ [1]

[1] In these axioms, conjunction takes precedence over implication. Variables occuring in the antecedents of implications are universally quantified, those occuring only in consequents are existentially quantified.

comes to realize that it must also have a back.

$$seat(x) \wedge back(y,x) \supset chair(x)$$

3. We extend our knowledge by learning further consequences in an implication. A child looks out the window for the first time at snow and sees that it is white,

   $$snow(x) \supset white(x)$$

   and then goes out and realizes it is also cold.

   $$snow(x) \supset white(x) \wedge cold(x)$$

4. We can increase the number of arguments that a predicate takes. A child realizes that *mother* is not a property, a predicate with one argument, $mother(x)$, but rather is a relation, a predicate with two arguments, $mother(x,y)$. A mother is not merely a mother, but a mother *of* someone.

5. We can introduce a new predicate as a specialization of an old predicate, as when we learn the concept of beagle as a specialization of the concept of dog.

   $$beagle(x) \supset dog(x)$$

   In fact, new predicates usually come in sets of two or more, as a result of recognizing a distinction in the general concept with some regularity and utility. When we have learned about beagles, we must also have learned that there are other kinds of dogs as well.

6. Combining operations 2 and 5, we can specialize the predicate in the antecedent of an implication. A child from its limited experience might believe that all dogs have floppy ears,

   $$dog(x) \supset floppy\text{-}ears(x)$$

   and then refine that knowledge to the belief that all beagles have floppy ears.

   $$beagle(x) \supset floppy\text{-}ears(x)$$

7. Combining operations 3 and 5, we can specialize the predicates in the consequent of an implication. A student may know that rabbits are mammals and rodents are mammals,

$$rabbit(x) \supset mammal(x)$$
$$rodent(x) \supset mammal(x)$$

   and then learn that rabbits are rodents.

$$rabbit(x) \supset rodent(x)$$

All of these incremental changes are made in order that the agent have a better theory of some aspect of the world. But that raises a central question—What do we mean by "better"? What makes a theory "good"? What are the criteria?

The answer for individual cognition is the same as for scientific theories. A good theory is one that explains a significant amount of the observable evidence with an economy of underlying principles—coverage and elegance. One theory can be better than another if it covers more of the observable evidence or if it explains the same evidence with fewer principles. When we learn a new particular fact or refine our knowledge by making the antecedents of implications more specific, we are increasing the coverage of the theory. It now leads to more correct and fewer incorrect conclusions. When we extend our knowledge by making the consequent of implications more specific we are also extending the coverage of the theory. Premises lead to more, or more specific, conclusions, and hence more can be explained.

These operations increase the coverage of a theory. But what makes a theory simpler, more economical, more elegant? We began to get a hint of it with Operation 7. Consider the case of the rabbit again. Suppose we know that a rabbit is a small mammal with large incisors and long ears.

$$rabbit(x) \supset mammal(x) \wedge small(x) \wedge large\text{-}incisors(x) \\ \wedge\, long\text{-}ears(x)$$

Suppose we also know that a rodent is a small mammal with large incisors.

$$rodent(x) \supset mammal(x) \wedge small(x) \wedge large\text{-}incisors(x)$$

When we learn that a rabbit is a rodent, we achieve a **compactification** of our theory. Now all we need to know about a rabbit is that it is a rodent with large ears.

$$rabbit(x) \supset rodent(x) \wedge long\text{-}ears(x)$$

The rest follows from its being a rodent.

This kind of modification to a knowledge base often has the character of a revelation. We feel that we have had a real insight and that we now see things more clearly. We are able to make numerous correct predictions whereas before we were often mystified at what occurred. Let me give an example from my own life. I have a colleague whom over the years I have had many good and even more bad interactions with, and I have observed him in interactions with others, noting many in which he was friendly and helpful and many others in which he was antagonistic. I was often puzzled at this. Then one time in a conversation about him with someone else, I had a sudden insight that made everything clear. I understood an underlying principle that explained nearly all these incidents. The underlying principle was—He was defending what he saw as his ecological niche. When he praised someone's work, it was either because he had been explicitly acknowledged as the leading researcher in that area on whose work future research should build, or because it had nothing to do with his own work. When he attacked someone's work it was because it addressed the same problems he had addressed but in a different way. I had known many particular facts about my colleague, but I had not discerned the pattern. I also knew about defending ecological niches and the causal consequences of that. Putting these two bits of knowledge together resulted in a significant compactification of my knowledge of my colleague. I could see his previous behavior as a deductive consequence of a powerful underlying principle. Moreover, I could now predict how he would react in new situations, where before I could not.

I had a similar sense of revelation with respect to Michel Foucault. I had read several of his books, had enjoyed them, found them insightful and inspiring, but I had no unified view of what he believed. I could not have predicted what he would say on some new topic. Then in a discussion group, someone made a random remark about Foucault that made everything fall into place for me. He referred to Foucault as having a "nostalgia for the good old days when the mad ran free". For me, that threw Foucault's work into a new light. It was an expression of an underlying principle that gave me new insight into what I had read before and enabled me to guess, at least in a general way, what he might say on other issues.

This notion of the individual agent attempting to construct better and better theories to explain the world casts the agent in the role of a kind of scientist, making, testing, and modifying hypotheses. Indeed, the characterization I have given of learning as theory change works as well whether the theory belongs to a computer, a human, a community of scholars, or a community of any sort. The way of knowing in each case may be different, and the ways of learning will be different in a corresponding fashion, but at a sufficiently abstract level, the same processes are occurring.

So we can see in the development of scientific knowledge that the great breakthroughs are events of a very similar sort. A sudden insight occurs, and a great compactification of the theory is achieved. What before was a large, heterogeneous collection of unrelated facts is now seen as a highly structured set of deductive consequences of a few underlying principles.

Galileo's insight that orbiting and falling are the same process is an example of this.

In geology, the theory of plate tectonics explained, in terms of the simple physical concept of plates sliding past, over and under each other, the various hitherto unconnected facts and phenomena of earthquakes, volcanos, uplift, the distribution of fossils, the structure of the ocean floor, the distribution of mountains, the shape of continents, and so on.

Analogy and metaphor work by very similar principles. Consider an example of Lakoff and Johnson's: "Marriage is a collaborative work of art." We know a great deal about marriage, about collaboration, and about the production of works of art. When we construct the metaphor, however, we are recognizing a subset of the properties of marriage and collaborative works of art that will support the metaphor. Then, in addition, the payoff of the metaphor, we see a number of other properties that we previously might not have attributed to marriage that we can carry over to marriage from collaborative works of art.

The general pattern for the experience is illustrated in Figure 1. Similari-

$P_1, P_2, P_3, Q_1, Q_2, Q_3 \quad | \quad P_1, P_2, P_3, R_1, R_2, R_3, R_4$

$\Rightarrow$

$Q_1, Q_2, Q_3 \qquad\qquad R_1, R_4$

P_1, P_2, P_3, R_2, R_3

Figure 1:

ties are recognized in two distant areas of one's knowledge and factored out to become an abstract domain on their own—the underlying pattern, a kind of more coarse-grained theory covering both phenomena (the P's in Figure 1). In addition, one of the initial domain may provide certain properties (R_2 and R_3) to the abstract domain that may then be inherited by the other initial domain. Knowledge has thereby been gained.

II.

What does all this have to do with poetry?

There is a common experience in the reading of poetry that bears some resemblance to this critical moment of insight in learning. It is characterized by a sudden clarity as to the meaning of the poem, where that meaning involves a new idea, a new connection between things. I find it easiest to explain in terms of the two distinct concepts of "coherence" and "cohesion".

Coherence is the underlying implicit unity of a text that makes it a single discourse rather than a concatenation of unrelated discourses. Cohesion is the surface manifestation of coherence. The coherence of a text, for example, may depend on its successive segments being similar exemplifications of the same general principle, whereas the cohesion is manifested in the coreference relations and parallel syntactic structure. Edward Gorey writes texts that are cohesive without being coherent (Gorey, 1972)

> It was already Thursday, but his lordship's artificial limb could not be found; therefore, having directed the servants to fill the baths, he seized the tongs, and set out at once for the edge of the lake.

By contrast, a common device that Ezra Pound uses is to write a coherent text in which cohesion is largely lacking (Pound, 1960, p. 46):

> And the quicker you go to the texts the less need there will be for your listening to me or to any other long-winded critic.
>
> A man who has climbed the Matterhorn may prefer Derbyshire to Switzerland, but he won't think the Peak is the highest mountain in Europe.

Now let us extend the notion of cohesion somewhat to include not only the **semantic** cohesive features of coreference, lexical resemblances, and so on, but also varieties of **sound** cohesion such as rhyme, alliteration, and a particular continuing meter. (Winfried Nöth (1993) refers to all of this as "endophoric iconicity".) It is perhaps one of the most salient features of most poems that they present us with a dense fabric of cohesion, where "cohesion" is understood in this broad sense.

Another very salient feature of most poems is that they are hard to read. It is not like reading prose. You have to work at it. What are we are doing when we are doing this work? We are attempting to come up with a coherent interpretation of the whole. Why do we keep at it? Why don't we simply dismiss the poem as a concatenation of fragments? I think it is precisely the dense fabric of cohesion that leads us on. With all this cohesion, the reader feels, there must be an underlying coherence.

The discovery of coherence then comes at a critical moment in the poem when we call forth and force into combination large, highly structured conceptual schemas that are normally distant in our knowledge base. Moreover, these schemas are frequently highly charged emotionally.

In brief, then, the critical moment happens when after a struggle toward coherence in a dense network of cohesion, we achieve coherence by bringing into combination heavily charged, hitherto unrelated areas of our knowledge base.

This characterization is similar to an account of the aesthetic experience proposed in psycholinguistics by Bever (1986). In his view an aesthetically satisfying experience is one that "stimulates a conflict in perceptual representations, which is resolved by accessing another representation that allows the two conflicting ones to coexist." (p. 316.) He has examined from this perspective such successful cultural artifacts as the song "Happy Birthday" and the rhythmic pattern "One shave and haircut, two bits." In his account of "Happy Birthday", for example, we cannot decide which of two possible keys the song is in until the ambiguity is resolved precisely on the mention of the person's name.

A fairly pure example of this in poetry can be seen in Ezra Pound's famous haiku, "In a Station of the Metro":

> The apparition of these faces in the crowd;
> Petals on a wet, black bough.

Here two powerful but unrelated images are presented to us individually and we are forced to discover their relation.

The learning schema that the poem is aimed at forcing is shown in Figure 2. We see two apparently unrelated statements juxtaposed in what is supposed to be a poem. We try to construct a coherence. Because the evidence is so sparse and the images so distant, there are many ways this

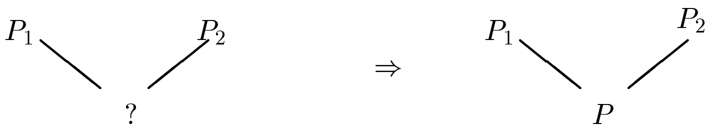

Figure 2:

could be done. Perhaps one way is to note a similarity of visual images and a sense of the anonymity of both the crowd and the petals.

Another fairly clear example of this phenomenon on a small scale is found in the Middle English poem,

> Western wind, when wilt thou blow?
> The small rain down can rain.
> Christ, that my love were in my arms,
> And I in my bed again.

The third line by itself creates an image of the couple embracing, standing up. The fourth line by itself creates an image of the poet lying in bed alone. It is when we try to put these two images together to form a coherent picture of the whole that we are forced to reinterpret them both as the couple lying in bed together, making love.

The experience that this poem promotes differs a bit from the Ezra Pound haiku. Each of the last two lines individually evokes an image that is more filled-out than is really justified by the text—the standing (P_3 in Figure 3) and the being alone (P_4). When we see the coherence between the two lines, we correct our image, and add a new property, making love (P_5).

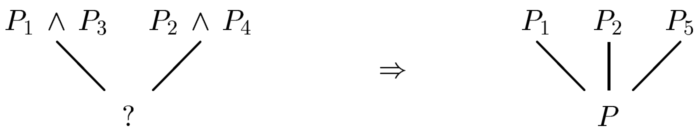

Figure 3:

There is a paradox here however. The world is very complex. As we go through Incremental Operations 1-7, our theories become more complex, they come to cover more of the observable evidence, and they are therefore better theories. But the operation of compactification makes the theory simpler. How is it that a simpler theory can be a better theory of a complex world? Especially since nearly every underlying general principle that yields a more compact theory, from Newton's laws to my general principles explaining Foucault and my colleague, are at best idealizations and at worst outright mistakes. Newton's laws don't work where there is friction or air resistance; plate tectonics enables only the coarsest sorts of predictions; there are numerous examples of my colleague not acting in defense of his ecological niche; and much in Foucault cannot be explained in terms of nostalgia for the good old days when the mad ran free.

The answer to this paradox involves something very fundamental about human cognition. We are finite creatures and we can only think a few things at once. One of the secrets to our success in the world is that we are able to view the world at different granularities, to discern different regularities at each of those granularities, to move among different granularities with remarkable ease, and to recognize when conclusions at one granularity hold or fail to hold at another.

We move from one granularity to another, coarser granularity by systematically excluding from consideration certain aspects or distinctions among things. When we view a highway as a line, we systematically ignore differences in width and surface roughness. When we talk about what Enlightenment philosophers believed, we systematically ignore the substantial differences between, say, Voltaire and Diderot, and focus on the similarities they share by virtue of their being Enlightenment philosophers. When we reason about the processes of plate tectonics, spans of centuries reduce to point-like entities. When I try to predict my colleague's behavior in terms of his defense of his ecological niche, I am systematically excluding the myriad of other factors that can determine events. When we deal with people solely as representatives of a particular occupation, discipline, nationality, or ethnic group, we are systematically excluding consideration of individual differences within groups, and drawing conclusions strictly on the basis of our theories of the behavior of "that sort of person", that is, our stereotypes.

Reasoning in coarse-grained theories is a very powerful cognitive tool. It allows us to draw quickly, from sparse evidence, broad-based conclusions that are usually correct.

But reasoning in a coarse-grained theory is only half of a functional cognitive process. The clue comes from a study a student of mine did a number of years ago. To determine how people describe shapes, he had a number of subjects describe electron micrographs. What the subjects typically did was describe the strange shapes they saw in terms of the closest prototypical shape, an ellipse, say, or a rectangular box. But then they described how the actual shape differed from the prototype, generally in terms of deformations that would have to be done on the prototype. For example, "a rectangular box where one of the corners has been stretched out to a sharp point". When we reason in coarse-grained theories, we often get close to the truth. But at that point we need to increase the granularity at which we are reasoning about the situation, to check and refine our conclusions in a more fine-grained way.

The power of simple, idealized or approximate theories therefore arises from their place in a two-step process:

1. We determine, using the idealized theory, the prototype that best approximates the actual situation.

2. We move to a finer-grained analysis to determine how the actual situation deviates from the prototype.

My theory of men in business suits is that they know their way around, so I might ask one for directions. But when I do, I will be attuned to signals that might indicate that he doesn't.

When we refine the granularity at which we view a domain, we are dividing the territory of that domain into a larger set of properties, each covering less territory. In Figure 4, we move from seeing the world in terms of coarse-grained predicates P and Q into a view of the world in terms of fine-grained predicates, the P_i's and Q_j's, which may not match P and Q exactly.

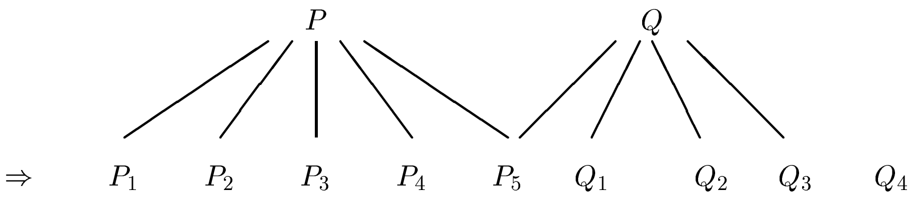

Figure 4:

It is important when we make theories more compact by recognizing underlying generalizations, not to lose the finer-grained view. It is moreover an important step in learning when we are able to devise a finer-grained theory of a domain, where before there was only a coarse-grained theory, especially if we also know how the theories at the two granularities articulate, so we can move easily between them.

IV.

Back to poetry.

Much of the best of literature invites us to a more fine-grained analysis of the issues in question. One such poem is this sonnet by Milton.

(1) Lawrence of virtuous father virtuous son,

(2) Now that the fields are dank, and ways are mire,

(3a) Where shall we sometimes meet, and by the fire
(4a) Help waste a sullen day;
(4b) what may be won
(5a) From the hard season gaining;
(5b) time will run
(6) On smoother, till Favonius reinspire
(7) The frozen earth; and clothe in fresh attire
(8) The lily and the rose, that neither sowed nor spun.

(9) What neat repast shall feast us, light and choice,
(10) Of Attic taste, with wine, whence we may rise
(11) To hear the lute well touched, or artful voice
(12) Warble immortal notes and Tuscan air?

(13) He who of these delights can judge, and spare
(14) To interpose them oft, is not unwise.

This poem is built around a highly charged tension, that between hard duty in the face of adversity and the easy comforts of relaxed fellowship. The tension operates in a much more interwoven fashion than it does in Pound's haiku or "Western Wind", but just as in those poems, juxtaposition seems to promise coherence and thus impels us to try to construct a coherence. In Milton's sonnet, we find that when we have constructed the coherence, we have done so by bringing together conflicting schemes for duty and delights, fitting them together into a more fine-grained analysis.

For example, there is a parallel relation between segment (3) - (4a) and (4b) - (5). We have two similar entities in the time periods, the day and the season, and their similarity is strengthened by their further similarly unpleasant properties: the day is sullen and the season is hard and gaining. This lexical similarity is a cohesive feature that leads us to expect parallel messages. But the property asserted of the day is that together we will waste it, whereas the property asserted (or questioned) of the season is that something will be won from it, presumably by us. To see these ideas as parallel and hence to see the text as coherent, we need to make an assumption that we would otherwise not necessarily make. We must see "waste together" and "win something from" as implying identical properties. We can do this by assuming that wasting time together is a good thing to do. This implicature turns out to be central to the meaning of the whole poem.

Between segments (2)-(5a) and (5b)-(8) there seems to be a causal relation. Wasting time together causes time to run on smoother. Note, however, that although the causal relation is strongly implied by the word "will", it is not explicitly signalled. To recognize the relation we have to assume it is one of the poet's beliefs that spending time together eases hardship. Thus, we are again forced to draw as an implicature a proposition that is very close to the central thrust of the poem.

In lines (9)-(12), nearly every phrase exhibits an interweaving of words of art and moderation with words of feasting, balancing enjoyment and control in a way that reflects in the small the balance between delight and duty urged by the poem as a whole. The "repast" is "neat", the "feast" is "light and choice", the "lute" is "well touched", and the "voice" is "artful". It is also significant that the cultures that are mentioned, "Attic" and "Tuscan", are cultures that are characterized by measured conduct. What is described is neither a Spartan repast nor a Roman feast.

In lines (13) - (14), the word "spare" by its own ambiguity suggests this balance so central to the meaning of the poem. The intended meaning of the word "spare" has been a matter of controversy throughout the critical history of the poem. One group of critics argues that the phrase "spare to interpose" is to be interpreted as "refrain from interposing", whereas the other group of critics argues that it means "spare time to interpose". The difference is, of course, complete. It is a question of whether the principal thrust of the poem is one thing or its opposite. I would argue, however, that the readings are not actually contradictory, but rather indicative of the moderation that is the pervading message of the entire sestet.

Let us examine our experience as we read carefully and attempt to construct coherent interpretations of successive initial segments of the poem. When all that has been seen is "... of these delights can judge, and spare", only one interpretation is possible. "Judge and spare" constitutes a conjoined verb phrase sharing the object "these delights". The sense must be "refrain from". This is reinforced by the adjectival meaning of "spare" as thin and lean, a counterpoint to the word "delights", not supported by the syntax, but perhaps primed in the reader's mind.

We continue on to the words "To interpose them". At this point, the "refrain from" meaning cannot possibly be correct. It would be in flat contradiction to the rest of the poem. It would perhaps be possible to resolve the contradiction, but the resulting interpretation—one must abstain from even moderate delights—would be quite jarring, in contrast to the easy, relaxed, and moderate tone of the entire sonnet. We are thus forced toward the other sense of "spare". We are being urged to spare time to interpose the delights, and this meaning meshes well with and in fact sums up the entire poem.

Then we come to the word "oft". Suddenly the "refrain from" sense becomes possible again, but this time in a way that does not contradict our previous interpretation. It is not interposing delights among our duties that we should refrain from, but doing so too often. We can take Milton to be saying that we should spare time to interpose delights, but we should refrain from interposing them too often. Both senses of "spare" can therefore be adopted, not in a contradictory but in a qualifying fashion.

In the ambiguity of the single word "spare" we thus see an example of something that is not uncommon in the best literature. The rhetorical device of paradox asserts something that is contradictory. In the rhetorical device of irony, a contradiction is implicit; what is said conflicts with what can be taken to be a belief shared by the speaker and hearer. We can take as the purest example of these devices the ordinary utterance, "Well, ... that's true and it isn't true." We do not want to convict the speaker of this utterance of inconsistency, and this forces us to reinterpret the utterance, or interpret it more deeply. The utterance is thus a way of saying, "The situation is not so simple." We interpret it by inferring something like "It is true in that P_1, P_2, and P_3, but it is not true in that Q_1, Q_2, and Q_3," and this elaboration is just the kind of more complex analysis of the situation that is being urged. The devices of paradox and irony are, in their highest uses, ways of conveying a complexity that is otherwise difficult to convey with the sometimes too blunt instrument of language. Milton has used the ambiguity of the word "spare" to the same effect. It is not the case simply that one should indulge in delights or that one should not indulge in delights. The situation is more complex. One must judge carefully and enjoy the delights in moderation.

For my final example, I want to look briefly at Winfried Nöth's (1993) analysis of William Blake's "Tyger, Tyger".

Tyger Tyger, burning bright,
In the forests of the night;
What immortal hand or eye,
Could frame thy fearful symmetry?

In what distant deeps or skies
Burnt the fire of thine eyes!
On what wings dare he aspire?
What the hand, dare seize the fire?

And what shoulder, and what art,
Could twist the sinews of they heart?
And when thy heart began to beat,
What dread hand? and what dread feet?

What the hammer? what the chain?
In what furnace was thy brain?
What the anvil? what dread grasp,
Dare its deadly terrors clasp?

When the stars threw down their spears
And water'd heaven with their tears:
Did he smile his work to see?
Did he who made the Lamb make thee?

Tyger Tyger, burning bright,
In the forests of the night;
What immortal hand or eye,
Dare frame thy fearful symmetry?

Nöth points out the manifold symmetries in this poem. The repetition of Stanzas 1 and 6, the fact that Stanzas 2 and 5 begin with descriptions of events in the skies and end with questions, and the fact Stanzas 3 and 4 each consist of five wh-questions, all create a kind of mirror symmetry between the first half and the last half. The "*a a b b*" rhyme pattern, moreover, constitutes a kind of translative symmetry.

But then he points out a number of instances of what he calls "symmetry breaking", points at which the close symmetry fails. Two of the most striking are in lines 4 and 24, the contrast between "Could" and "Dare", and the fact that the only blemish in the perfect rhyme of the poem is between "eye" and "symmetry". Nöth argues that works of art must not suffer from an excess of symmetry, and that this work of art is, in a sense, about this very fear of symmetry.

This “symmetry breaking” in the otherwise smooth pattern of the poem is a way for the poet, at the same time as he or she is proposing a particular view of the world, to indicate that the world is too complex to yield to any particular pattern. The poem proposes a compact, coarse-grained theory, the underlying pattern according to which the poem is organized, but at the same time it undercuts itself by indicating that the world inevitably escapes theories, and at the very least we must examine things at a finer granularity as well.

V.

Horace’s characterization of the function of literature is that it is to “instruct and delight”. In this talk I have focused on the “instruction”. I have viewed the poet as someone who is attempting to induce a critical change in our knowledge, either by inducing us to see similarities where before we did not, resulting in a compactification of our knowledge and its resulting clarity, or by inducing us to examine phenomena more deeply, at a finer granularity.

But what about delight? Bever (1986) has a partial answer for this. He locates the aesthetic experience in the joy we have in solving problems. Human beings are goal-directed, planning agents. We have desires that the world does not yield to easily. To achieve our goals, we have to understand the world in order to be able to manipulate it. Any breakthrough in making sense of our experience is at the same time a breakthrough in our ability to achieve what we desire, and should give us a corresponding joy.

Acknowledgements

This research is based in part on work supported by the National Science Foundation under Grant IRI-9304961 (Integrated Techniques for Generation and Interpretation). Any opinions, findings, and conclusions or recommendations expressed in this article are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

[1] Bever, Thomas G., 1986. “The Aesthetic Basis for Cognitive Structures”, in Myles Brand and Robert M. Harnish (Eds.), *The Representation of*

Knowledge and Belief, pp. 314-355, The University of Arizona Press, Tucson, Arizona.

[2] Gorey, Edward, 1972. *Amphigorey*, Longmans Canada Ltd., Toronto, Canada.

[3] Hobbs, Jerry R., Mark Stickel, Douglas Appelt, and Paul Martin, 1993. "Interpretation as Abduction", *Artificial Intelligence*, Vol. 63, Nos. 1-2, pp. 69-142.

[4] Nöth, Winfried, 1993. "Cognition, Iconicity, and Blake's Fearful Symmetry", paper presented at the Symposium on Literary Cognition, Information, Communication: Theses and Prospects, Carleton University, Ottawa, Canada, June 1993.

[5] Peirce, Charles Sanders, 1955. "Abduction and Induction", in Justus Buchler, editor, *Philosophical Writings of Pierce*, pp. 150-156, Dover Books, New York.

[6] Pound, Ezra, 1960. *ABC of Reading*, New Directions, New York, NY.