“Pulling it All Together” via Psychometric AI

Selmer Bringsjord and **Bettina Schimanski**
selmer@rpi.edu • schimb@rpi.edu
Department of Cognitive Science
Department of Computer Science
Rensselaer AI & Reasoning (RAIR) Lab
Rensselaer Polytechnic Institute (RPI)
Troy NY 12180 USA

Abstract

Rather long ago, Newell (1973) wrote a prophetic paper that could serve as a rallying cry for this 2004 symposium: “You Can’t Play 20 Questions with Nature and Win.” A number of those concerned with integrated cognition know of this paper, which helped catalyze both modern-day computational cognitive modeling through cognitive architectures (like ACT-R, Soar, Polyscheme, CLARION, etc.), and AI’s attempt to build a chess-playing machine better at the game than any human. However, not many know that in this paper Newell suggested a third avenue for achieving integration, one closely aligned with psychometrics. In the early days of AI, at least one thinker started down this road for a time (Evans 1968), but now the approach is long dead and all but forgotten. We recommend resurrecting this approach, in the form of what we call *Psychometric AI*, or just PAI (pronounced to rhyme with “π”). We briefly describe and defend PAI herein. We include some coverage of PERI, a robot in our lab who exemplifies PAI and integrated (artificial) cognition. We also explain how it is that we can strive for integration via the near-exclusive use of mechanized logic, under the umbrella of Newell’s third route.

Introduction

Rather long ago, Newell (1973) wrote a prophetic paper that could serve as a rallying cry for this 2004 symposium: “You Can’t Play 20 Questions with Nature and Win.” A number of those concerned with integrated cognition know of this paper, which helped catalyze both modern-day computational cognitive modeling through cognitive architectures (like ACT-R, Soar, Polyscheme, CLARION, etc.), and AI’s attempt to build a chess-playing machine better at the game than any human. However, not many know that in this paper Newell suggested a *third* avenue for achieving integration, one closely aligned with psychometrics. In the early days of AI, at least one thinker started down this road for a time (Evans 1968), but now the approach is long dead and all but forgotten. We recommend resurrecting this approach, in the form of what we call *Psychometric AI*, or just PAI (rhymes with “π”). We briefly describe and defend PAI herein. Specifically, our plan is as follows. In Section 2 we briefly present Newell’s call for (as we see it) PAI in his seminal “20 Questions” paper. Section 3 provides a naive but serviceable-for-present-purposes definition of PAI in line with Newell’s call. Section 4 provides some information about our main current implementation of PAI in the form of the robot PERI. We briefly describe the “building block” approach that flows from PAI in Section 5, and in Section 6, we present and rebut some objections to PAI.[1] It is revealed in this section that though our research is logic-based, if Newell’s third avenue is sound, an ecumenical approach isn’t needed for achieving an integrated system. The final section points specifically to what we’re currently working on within PAI.

Newell and the Neglected Route Toward Integration

In his “20 Questions” paper, Newell (1973) bemoans the fact that, at a symposium gathering together many of the greatest psychologists at the time, there is nothing whatsoever to indicate that any of their work is integrated. Instead, everyone is carrying out work (of the highest quality, Newell cheerfully admits) on his or her own specific little part of human cognition. In short, there is nothing that, to use Newell’s phrase, “pulls it all together.” He says: “We never seem in the experimental literature to put the results of all the experiments together.” (Newell 1973, p. 298) (It seems to us that the same observation could be made today about systems engineered in AI.) After making clear that he presupposes that “man is an information processor,” Newell offers three possibilities for addressing the fragmentary nature of the study of mind.

The first possibility he calls “Complete Processing Models.” He cites his own work (with others; e.g., Simon) based on production systems, but makes it clear that the production system approach isn’t the only way to go. Of course today’s cognitive architectures [e.g., SOAR (Rosenbloom, Laird, & Newell 1993); ACT-R (Anderson 1993; Anderson & Lebiere 1998; 2003); CLARION (Sun 2001); and Polyscheme (Cassimatis 2002; Cassimatis *et al.* 2004)] can be traced back to this first possibility.

The second possibility is to “Analyze a Complex Task.” Newell writes:

[1] A more robust list of objections/rebuttals can be found in (Bringsjord & Schimanski 2003).

> A second experimental strategy, or paradigm, to help overcome the difficulties enumerated earlier is to accept a single complex task and do all of it . . . the aim being to demonstrate that one has a significant theory of a genuine slab of human behavior. . . . A final example [of such an approach] would be to take chess as the target super-task. (Newell 1973, p. 3003–304)

This second possibility is one most people in cognitive science and AI are familiar with. Though Deep Blue's reliance upon standard search techniques having little cognitive plausibility perhaps signaled the death of the second avenue, there is no question that, at least for a period of time, many researchers were going down it:

> Early chess systems sought to duplicate or mimic the methods of humans. But this proved to be far too difficult: What precisely suggests any particular move? Instead, successful chess programs capitalize on the particular strengths of computers: rapid and massive parallel search. (Stork 1997)

The third possibility, "One Program for Many Tasks," is the one people seem to have either forgotten or ignored. Newell described it this way:

> The third alternative paradigm I have in mind is to stay with the diverse collection of small experimental tasks, as now, but to construct a single system to perform them all. This single system (this model of the human information processor) would have to take the instructions for each, as well as carry out the task. For it must truly be a single system in order to provide the integration we seek. (Newell 1973, p. 305)

For us, the system in question is the robot PERI, intended to fit into a specific mold within Newell's third possibility, viz.,

> A . . . mold for such a task is to construct a single program that would take a standard intelligence test, say the WAIS or the Stanford-Binet. (Newell 1973, p. 305)

PERI is for us an implementation of PAI, and to an explication of this brand of AI we now turn.

What is Psychometric AI?

What is AI? We'd be willing to wager that many of you have been asked this question — by colleagues, reporters, friends and family, and others. Even if by some fluke you've dodged the question, perhaps you've asked it yourself, maybe even perhaps (in secret moments, if you're a practitioner) *to* yourself, without an immediate answer coming to mind. At any rate, AI *itself* repeatedly asks the question — as the first chapter of many AI textbooks reveals. Unfortunately, many of the answers given don't ensure that AI tackles head on the problem of integrated cognition.[2] Our answer, however, is one in line with Newell's third possibility, and one in line with a perfectly straightforward response to the "What is AI?" question.

To move toward our answer, note first that presumably the 'A' part of 'AI' isn't the challenge: We seem to have a fairly good handle on what it means to say that something is an artifact, or artificial.[3] It's the 'I' part that seems to throw us for a bit of a loop. What's intelligence? *This* is the big, and hard, question. Innumerable answers have been given, but most thinkers seem to forget that there is a particularly clear and straightforward answer available, courtesy of the field that has long sought to operationalize the concept in question; that field is psychometrics. Psychometrics is devoted to systematically measuring psychological properties, usually via tests. These properties include the one most important in the present context: intelligence. In a nutshell, then, the initial version of our account of intelligence is this: *Some agent is intelligent if and only if it excels at all established, validated tests of intelligence*.

We anticipate that some will insist that intelligence tests, even broad ones, are still just too narrow, when put in the context of the full array of cognitive capacities seen in *homo sapiens*. Well, we agree! But we are understanding intelligence, from the standpoint of psychometrics, to include many varied tests of intellectual ability. Accordingly, we actually work on the basis of a less naive definition of PAI:

> Psychometric AI is the field devoted to building information-processing entities capable of at least solid performance on all established, validated tests of intelligence and mental ability, a class of tests that includes not just the rather restrictive IQ tests, but also tests of artistic and literary creativity, mechanical ability, and so on.

This definition, when referring to tests of mental ability, is referring to much more than IQ tests. For example, following (Sternberg 1988), someone with much musical aptitude would count as brilliant even if their scores on tests of "academic" aptitude (e.g., on the SAT, GRE, LSAT, etc.) were low. But specifically what sorts of additional tests would be involved? We don't have space to canvass the myriad tests that psychometricians have validated. To give a quick sense of how latitudinarian (and therefore challenging) Psychometric AI is intended to be, we mention The Torrance Tests of Creative Thinking (Torrance 1990; 1988). This test comes in both "visual" and "verbal" forms. In the visual form, test takers are asked to draw pictures (often by enriching existing sketches); in the verbal form, test takers are asked to write — creatively. For example, one of the activities subjects engage in on the verbal test is:

> Most people throw their paper towel rolls away, but they have thousands of interesting and unusual uses. In the spaces below and on the next page, list as many of these interesting and unusual uses as you can think of. Do not limit yourself to any one size of roll. You may use as many as you like. Do not limit yourself to the uses you have seen or heard about; think about as many possible new uses as you can. (*Similar* to the verbal version of (Torrance 1990).)

AI then reduces to Psychometric AI: the field devoted to building a computational system able to score well on such

[2] E.g., Russell & Norvig (2003) seem to us to characterize AI in a way (via functions from percepts to actions; they call these functions *intelligent agents*) that fosters a focus on *non*-integrated cognition.

[3] We can ignore here conundrums arising from self-reproducing systems, systems that evolve without human oversight, etc.

tests. This may strike you as a preposterously narrow definition of AI. The first step in diffusing this attitude is to take a look at some intelligence tests, some of which, we surmise, are a good deal richer than you might at present think. In short, in choosing the WAIS (Wechsler Adult Intelligence Scale)[4], Newell knew what he was doing.

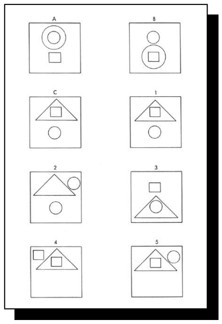

Figure 1: Sample Problem Solved by Evan's (1968) ANALOGY Program. Given sample geometric configurations in blocks A, B, and C, choose one of the remaining five possible configurations that completes the relationship: *A is to B as C is to . . .?*

In the early days of AI, Psychometric AI was at least implicitly entertained. After all, in the mid 1960s, the largest Lisp program on earth was Evans' (1968) ANALOGY program, which could solve problems like those shown in Figure 1. Evans himself predicted that systems able to solve such problems would "be of great practical importance in the near future," and he pointed out that performance on such tests is often regarded to be the "touchstone" of human intelligence. Unfortunately, ANALOGY simply hasn't turned out to be the first system in a longstanding, comprehensive research program (Newellian or otherwise): after all, we find ourselves, at present, trying to start that very program. What went wrong? Well, certainly Psychometric AI would be patently untenable if the tests upon which it is based consist solely of geometric analogies. This point is entailed by such observations as this one from Fischler & Firschein (1987):

> If one were offered a machine purported to be intelligent, what would be an appropriate method of evaluating this claim? The most obvious approach might be to give the machine an IQ test. . . . However, [good performance on tasks seen in IQ tests would not] be completely satisfactory because the machine would have to be specially prepared for any specific task that it was asked to perform. The task could not be described to the machine in a normal conversation (verbal or written) if the specific nature of the task was not already programmed into the machine. Such considerations led many people to believe that the ability to communicate freely using some form of natural language is an essential attribute of an intelligent entity. (Fischler & Firschein 1987, p. 12)

Unfortunately, while this quote helps explain why ANALOGY in and of itself didn't ignite a research program to drive AI, Fischler & Firschein apparently are familiar with only what we call **narrow**, as opposed to **broad**, intelligence tests — and Newell, referring to the WAIS, had in mind the broad ones. Arguably, this distinction goes back to Descartes' (1911, p. 116) claim that while a machine could in the future pass any test for a particular mental power (including, before Turing was born, the test that now bears his name), no machine could pass a test for any mental power whatsoever. This rather speculative claim can perhaps be cashed out in two different and longstanding views of intelligence within psychology: Thurstone's (1938) and Spearman's (1927). In Thurstone's view (put barbarically), intelligence consists in the capacity to solve a *broad* range of problems, e.g., verbal analogies, geometric analogies, digit recall, story understanding, commonsense reasoning, arithmetical calculation, and so on. In Spearman's view (again, put roughly), intelligence is a specific, *narrow*, underlying capacity (notoriously) referred to as g, summoned up to the highest degree when solving highly focused and abstract problems like those ANALOGY solved. The most famous set of "g-relevant" problems is the tightly guarded and much-used Raven's (1962) Progressive Matrices, or just 'RPM.' An example of a problem from RPM is shown in Figure 2, which is taken from (Carpenter, Just, & Shell 1990). As part of the PERI project, we have built theorem prover-based agents able to infallibly crack not only geometric analogies, but RPM items they have never seen before (Figure 2 shows part of an OTTER (Wos *et al.* 1992) proof that serves to identify the solution). (The algorithms deployed by these agents were devised as part of contracted research for the Educational Testing Service, or ETS.) It is much harder to build agents able to solve broad tests of intelligence — tests that include sub-tasks demanding the kinds of communicative capacities Fischler & Firschein have in mind.

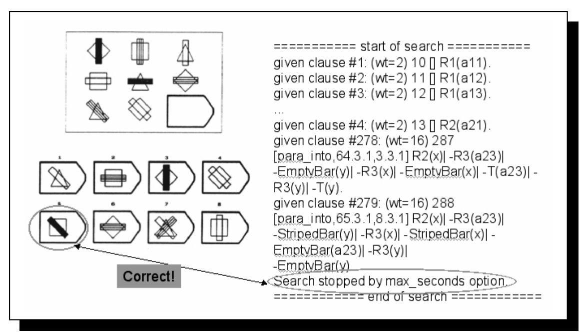

Figure 2: Simple RPM Problem "Cracked" by RAIR Lab's PERI

Psychological Corporation's popular WAIS (Wechsler Adult Intelligent Scale), the very test Newell cited, is a paradigmatic example of a broad intelligence test that includes the full array of "Thurstonean" sub-tests (the complete array is enumerated in Baron (2000)). Tables 1 and 2, taken from (Baron & Kalsher 2000), summarize the wide array of tasks on the WAIS. It should be clear that when Newell described the third possibility as one program for

[4]'Wechsler Adult Intelligence Scale' and 'WAIS' are trademarks of Harcourt Assessment, Inc., registered in the United States of America and/or other jurisdictions.

Subtest	Brief Description
Picture Completion	Examinees indicate what part of each picture is missing.
Picture Arrangement	Examinees arrange pictures to make a sensible story.
Block Design	Examinees attempt to duplicate designs made with red and white blocks.
Object Assembly	Examinees attempt to solve picture puzzles.
Digit Symbol	Examinees fill in small boxes with coded symbols corresponding to a number above each box.

Table 1: Performance Subtests of the WAIS
(Baron & Kalsher 2000)

Subtest	Brief Description
Information	Examinees are asked to answer general information questions, increasing in difficulty.
Digit Span	Examinees are asked to repeat series of digits read out loud by the examiner.
Vocabulary	Examinees are asked to define thirty-five words.
Arithmetic	Examinees are asked to solve arithmetic problems.
Comprehension	Examinees are asked to answer questions requiring detailed answers which indicate their comprehension of the questions.
Similarities	Examinees indicate in what way two items are alike.

Table 2: Verbal Subtests of the WAIS
(Baron & Kalsher 2000)

many tasks, and pointed at the WAIS, he was making good sense.[5]

With help from additional researchers in the RAIR Lab, we are in the process of "cracking" the WAIS, by way of the design and construction of PERI. The sub-test we have cracked most recently is "Block Design." PERI, when given any configuration of blocks in the space of all possible ones, reasons to the solution in under a second of CPU time, and proceeds to assemble this solution with its manipulator/gripper. In the next section we provide some details about some of PERI's exploits.

The Robot PERI: Newell's "One Program"

PERI (Psychometric Experimental Robotic Intelligence) is (on its way to being) an integrated system capable of logic/reasoning, vision, physical manipulation, speech, and hearing. It interacts with the environment via its five-degree-of-freedom vertically articulated mechanical arm, a SCORBOT-ER IX model from the Intelitek Corporation. It's original hand was a pneumatic two-finger parallel gripper which could either be completely opened or closed around an object. This has recently been upgraded to a more complicated but agile BH8-260 BarrettHand Dexterous 3-Finger Grasper System. The fingers can be rotated about a center

[5]It's clear as well that Fischler & Firschein's criticism of simplistic versions of Psychometric AI certainly evaporates in the face of the WAIS. That this is so follows from the sub-test on the WAIS known as "Comprehension," in which, in ordinary conversation, subjects are asked fiendishly tricky "general knowledge" questions. For example, examinees might be asked to explain why copper is often used in making electrical wires or what the advantage is of keeping money in a bank account.

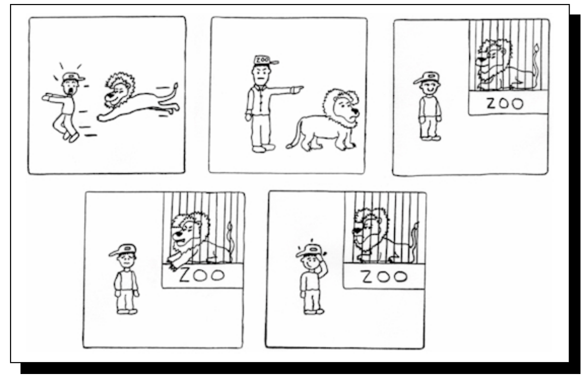

Figure 3: This is a home-grown example of the Picture Arrangement subtask from the WAIS (but not an actual sample from it). *Examinees must arrange the snapshots to make a coherent story.*

base (like a palm) to work oppositionally or together simultaneously. PERI's range of motion is still limited compared to a human arm/hand, but for its current purposes it does fairly well. Its vision is based on the output of a Sony Black-and-White XC55 Video Camera and Cognex MVS-8100M frame grabber. PERI's speech is transmitted through computer speakers and it hears through a microphone attached to the speaker's head while using the Dragon Naturally-Speaking Professional Solutions Version 6.0 software. At the core of PERI resides its brain and nervous system — a complex Common Lisp program and an associated Scorbot Advanced Control Language Library. Due to the lack of space, this is as much technical detail as we can communicate about PERI.

The rest of this section will focus on the Block Design task and PERI's success with it. PERI can not only solve the particular Block Design problems in the WAIS, but any Block Design puzzle expressible in first-order logic. For legal reasons we are unable to disclose the actual Block Design task from the WAIS, therefore we discuss another similar yet even more challenging block puzzle (courtesy of the Binary Arts Corporation). (In the interests of space, we leave aside the mathematization of the WAIS block puzzle and the harder one from Binary Arts. For ease of exposition, we refer to the space in question as $\mathcal{S}$.)

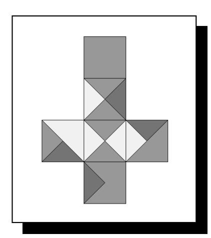

Figure 4: One Puzzle Block Folded Out

In this particular puzzle there are a total of four blocks, each of which is different. There are only three colors (pink, purple, and yellow) used to make the design on each side;

that design is either a combination of up to four triangles or one solid color. This is done merely to give a specific color to each edge of a block. In fact, all the sides of all four cubes are different from one another. This means there are a total of 24 unique sides. Refer to Figures 4 and 5 for a closer look.

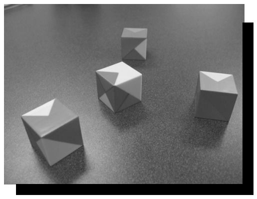

Figure 5: Blocks (from Binary Arts' # 5766) Scattered

The task is (after having been presented with the cubes for the first time) to place them together so that every edge that touches another is the same color. All cubes must be used, but obviously there are quite a few different solutions. One solution is shown in Figure 6. Does the task sound easy to you? If so, you are supremely confident. While PERI solves the hardest configuration in a matter of seconds, after having visually examined the blocks, in our experience it can take a clever human several long minutes, if not a half hour, to conquer the entire task. Figure 7 shows PERI assembling a solution.

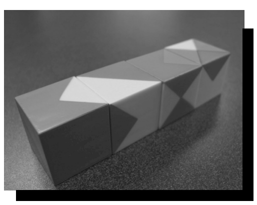

Figure 6: A Solution to Binary Arts Corp.'s Puzzle 5766

The same basic algorithm (described below) that PERI uses for the WAIS Block Design task can be used to solve any such puzzle in the overall mathematical space $\mathcal{S}$ ($\approx$ 3D regular solid with each side having a characteristic capturable in extensional first-order fashion). The first step is to encode the pieces as declarative information about geometric shapes in PERI's "mind." Before the test is administered to a human participant, he is given a chance to examine the blocks (or other shapes), as would PERI. What follows is a general algorithm which PERI can apply to any 3D physical shapes within a limit of size (i.e., which its gripper can properly hold and manipulate).

General Algorithm for "Cracking" 3D Block Design in $\mathcal{S}$

1. Document original pieces by color, dimension, characteristics on each side, and total number.
2. Input goal configuration (a picture that will need to be deciphered).
3. Partition the goal into distinguishable pieces that match similar aspects of those that are available pieces in the original. Start first with the entire goal as one piece. Some aspects of the pieces may be ignored at this stage.
4. Once the goal has been partitioned, determine if original puzzle pieces match the partitioned ones. If not, go back to step 3 and partition it into two pieces, three pieces, etc. (An exceedingly large cutoff is imposed to handle cases where no partitioning is valid, otherwise non-halting is possible.) If there are matching original puzzle pieces to the goal partitioning, go on to step 5.
5. Start with a goal piece and match it to an original piece that has not yet been used. There will be a finite search for each matching piece since step 4 has been passed, indicating the goal is known to be solvable. When a match is found, the original piece is physically added to the solution "arena" by changing the $< x, y >$ positioning of the original piece as well as the angle, side, or any other necessary aspect. Continue the present step until no more pieces in the goal exist that need a match.

In the case of the puzzle from the Binary Arts Corporation, the goal configuration is not specified ahead of time (as it is in the WAIS). Therefore, we assume that the goal is given ahead of time and are then able to use the above general algorithm without any modification. PERI can solve the original version of Binary Arts' puzzle; however, the original version doesn't correspond to the WAIS Block Design task, the cracking of which was our goal. The next challenge PERI faces is the more difficult "Picture Arrangement" subtask of the WAIS (see Figure 3), and our presentation will include a report on our progress with this subtask, which is briefly discussed in the final section.

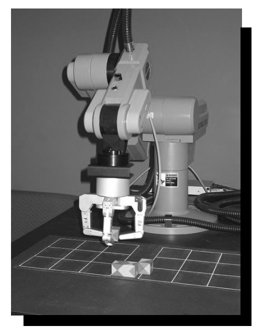

Figure 7: PERI Solving a Block Design Puzzle (with new three-fingered hand)

The PAI-Based "Building Block" Approach

With Newell, our goal is integrated intelligence via his "third paradigm." However, we believe that PAI is a fertile approach from an *engineering* standpoint. We believe this because the WAIS's tasks, once mechanized, will provide us with "building blocks" from which we can create larger, more complex systems, which will themselves be integrated. If we're right, PAI may be *doubly* integrative.

Figure 8 conveys pictorially the idea of using different subtests from certain chosen tests of mental and/or physical ability to create a composite artificial agent for a specific application.

Assume that our lab manages to "crack" a slew of intelligence tests in pursuit of Newell's dream; assume, specifically, that we crack: RPM, the WAIS, TTCT (Torrance Tests of Creative Thinking; see footnote) *and* PDMS-2 (Peabody Developmental Motor Scales-Second Edition), a test of early childhood motor development skills (both fine and gross). Under these assumptions, what sort of robots might we find

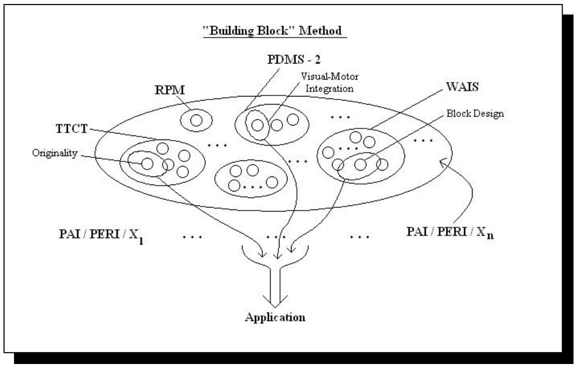

Figure 8: This diagram provides a pictorial view of the Building Block Method, depicted here with the following tests: Torrance Tests of Creative Thinking (TTCT), Raven's Progressive Matrices (RPM), Peabody Developmental Motor Scales-Second Edition (PDMS-2), and the Wechsler Adult Intelligent Scale (WAIS)

ourselves poised to build? Well, we confess that we don't exactly know. We have a lot of PAI-based R&D to carry out before we're in this "catbird seat." However, it's probably safe to say that these building blocks would enable us to engineer an artificial waiter — a robot able to carry food from kitchen to table, take orders from customers, and discuss with them items on the menu and topics that may come up in ordinary conversation. In this case, creativity will likely be needed for conversation in general, as well as for handling dynamic situations that need potentially unique solutions, such as a customer with a request for a meal that is not offered on the menu. Algorithms used to solve the TTCT would be useful here, as would the algorithm for the Comprehension subtask of the WAIS. In addition, of course, much mechanical ability is needed (to avoid obstacles (e.g., glasses and other items already on the table), hand out and retrieving menus, and carry the food to the table as well as clear when done).

Some Objections

We now present and rebut four objections to PAI.

Objection 1: "How can Newell's third possibility be on the right track? After all, PAI aimed at the WAIS is so idiosyncratic!"

Actually, Psychometric AI is far from idiosyncratic, because it is (at least arguably) a generalization of a longstanding answer to the "What is AI?" question, namely, the answer that appeals to the Turing Test (TT) and its relatives. To see this answer in action, let's turn to *AIMA* (Russell & Norvig 1994), which tells us that there are four general, different ways to define intelligence (pp. 4–8): we can say that an entity is intelligent if and only if it "thinks like humans", "acts like humans", "thinks rationally", or "acts rationally."

Russell & Norvig (1994) opt for the fourth route, but we want to draw your attention to the first and third ones, which don't seem exactly promising, because 'thinking' is probably no clearer than 'intelligence.' However, Turing came up with the test that now bears his name precisely because he found the concept of thinking hopelessly vague. As Russell and Norvig point out, TT and other more stringent tests, e.g., Stevan Harnad's (1991) *Total* Turing Test (in which a passing robot must display not only human-level linguistic performance, but sensorimotor performance at this level as well), provide a way to clarify the first route in the quartet. Specifically, it can be said that AI is the field devoted to building artificial entities (or systems) able to pass TTT.

We could go on to present case after case in which TT or variants are used to define AI (e.g., see Ginsberg's (1993) introductory AI text), but perhaps you will agree that whether or not you affirm the TT-based answer to the "What is AI?" question, you have to admit that Turing, by the lights of a good many, is definitely on to something. But what, exactly? Well, no doubt there is more than one reason for the apparent immortality of the TT. But surely one reason for its longevity is simply that it's a clean, crisp *test*. Tests are attractive in no small part because they have determinate starts and ends, and yield concrete verdicts that can silence unproductive debate.[6] Turing's (1950) goal in his seminal "Computing Machinery and Intelligence" was indeed to supplant the maddeningly vague "Can a computing machine think?" with "Can a computing machine pass the Turing test?" We're not concerned here with whether he reached his goal. Rather, the idea is that PAI extends and clarifies Turing's approach, and that Newell, in pointing toward the WAIS, found this general approach to be promising.

Objection 2: "But Don't TT and TTT Subsume Psychometric AI? Why couldn't Newell simply have called for one system capable of passing TT/TTT?"

We offer a three-part rebuttal: (1) In an attempt to build "one system for many tasks," certainly "divide and conquer" is a prudent strategy, and Psychometric AI automatically enforces that methodology: tests are crafted to check for different capacities in focused fashion. Since all topics are fair game in TT and TTT, they have much less value as engineering goals than the WAIS. (2) There is another reason why PAI can be viewed as a helpful generalization of the Turing/Harnad approach. This reason is that, let's face it, neither TT nor TTT is currently a meaningful objective: they are both gigantically ambitious goals, so much so that no one has really made any progress toward reaching them.[7] Psychometrics offers us an abundance of tests, many of which are *reasonable* challenges for an artificial agent. Psychometric AI appropriates these tests. (3) The tests in question haven't been picked out of thin air. These tests allow us to look into the heart of mind/brain. That's the beauty and power of tests, when they have been empirically and statistically validated. Tests have a gem-like character, and PAI

[6] Of course, computability theory relies heavily on tests. E.g., when we say that a set A is **decidable**, we are among other things saying that we can successfully apply a test to some object o in order to determine whether or not $o \in A$.

[7] At *Turing 2000*, the conference held at Dartmouth to commemorate both Turing's 1950 prediction that a TT-passing machine would be created before the new millennium and AI's inaugural 1956 conference at Dartmouth, no system better at conversation than a toddler appeared.

piggybacks on this. Given this, if we build an agent able to pass established tests, we can be fairly confident that as a welcome side-effect we will have an agent capable of many significant higher-level feats.

Objection 3: "But AI has applications that need to be built!"

Of course. And Newell wasn't saying that *all* work should focus on building machines able to succeed on the WAIS. After all, few work directly on building an agent able to pass such tests. The idea was that an agent able to excel on the WAIS will have an ensemble of integrated powers — and that full package might well enable the construction of many desired applications. Recall our discussion of the "building block" nature of PAI.

Objection 4: "But your lab predominantly pursues logic-based AI. Therefore PAI targeted at the WAIS will be severely limited; in fact, integration itself will be precluded."

Recall that Newell enumerated three *distinct* responses to the fragmentary nature of the science of intelligence. The first response, "Complete Processing Models," is not the same as the third, our choice, "One Program for Many Tasks" (where those tasks come from the WAIS). While it is surely open to us and others to pursue this program (i.e., PERI) by harnessing *both* logicist and sub-symbolicist techniques, we do plead guilty: our work is indeed based mostly on logical systems.[8] (Though we didn't provide details given space constraints, Block Design makes use of the situation calculus, and hence the system $\mathcal{L}_I$.) But note that that, by Newell's own lights, is fine. The reason is that if it *isn't* fine, if the the underlying approach to taking the third avenue must itself accord with "Complete Processing Models," then Newell's third possibility would collapse into the first! Since we wholeheartedly subscribe to Newell's third route, and that route is distinct from the other two, there is nothing *in principle* wrong with our allegiance to logic-based AI [(Bringsjord & Ferrucci 1998; Genesereth & Nilsson 1987; Nilsson 1991; Bringsjord, Arkoudas, & Schimanski forthcoming)] for this particular project. Of course, even "Block Design" on the WAIS requires robotic manipulation, and therefore cognitive processing that is rapid and reactive. This may cause some to say that in *practice* our logicist method will fail. We don't think so (since, among other reasons, the RAIR Lab has some mobile robots that negotiate their environments solely on the basis of mechanized logic), but time and time alone will tell.

Current State of PAI: Two New Challenges

At present we are tackling two additional subtasks on the WAIS: "Picture Arrangement" and "Object Assembly".

Picture Arrangement

Picture Arrangement requires of examinees that they arrange jumbled snapshots to form coherent stories. (A home-grown example is shown in Figure 3. For legal reasons, actual WAIS examples cannot be shown in print.) Here our research attempts to make use of prior work in story generation (e.g., Bringsjord & Ferrucci 2000). Our attempt to enable PERI to crack "Picture Arrangement" is the first step in the exploration of this new dimension. In our initial work on this problem, once a snapshot in a group like that shown in Figure 3 is selected, a search for a consistency proof that a particular successor is possible under narrative constraints is fired. If a proof is found, the successor is selected, and the process iterates. Narrative constraints are declarative formalizations of plots, themes, and characters.

[8]In a sense derived from 'logical system' as featured in Linström's Theorems. See (Ebbinghaus, Flum, & Thomas 1984).

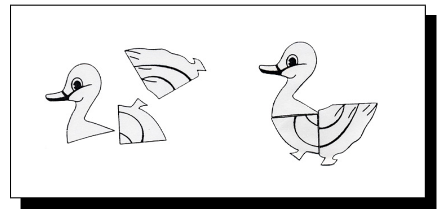

Figure 9: Sample Object Assembly item from the WAIS as provided by the publisher.[9] Given the scattered pieces on the right, the examinee must construct a "valid" image such as the finished puzzle shown on the left.

Object Assembly

The examinee taking the Object Assembly task is presented with several two-dimensional puzzle-like pieces laid out on the table and is then simply asked to put them together to make something (see Figure 9). No more detailed information is given. It is possible that most humans would spend a considerable amount of time first trying to figure out what the final shape/object will be. This may or may not help an individual put the pieces together with more efficiency (and may even waste time); however, it is within reason to assume that such a feat is undoubtedly difficult with, say, a hundred puzzle pieces, in which case another tactic (most likely trial-and-error) must be employed. Below we give a sketch of a general algorithm that could be applied to such a test for any 2D puzzle:

Proposed General Algorithm Sketch for "Cracking" 2D Object Assembly

1. Document original pieces by color, dimension, and any other unambiguous characteristics.
2. (First) attempt lining up only straight-lined sides, ignoring all irregular sides.
3. Choose (one of) the "largest" piece(s) to start with.
4. Choose the next largest piece (may be approximate/same size as first).
5. In a finite search, line up all straight edges together, only claiming a "match for a "perfect fit (dimensions must be "close enough). (This includes pieces that fit entirely within other pieces, or those in which at least two adjacent straight edges of one piece line up with another).
6. If no sides fit, cast this recently choosen piece aside and move on to the next largest piece.
7. Continue until either a piece is found that matches the first or none of them seem to match.

(a) If a match is found, these two pieces are now considered one large piece (i.e. no longer two separate pieces). Repeat algorithm starting at step 3 using this combined piece as the starting piece.

(b) If no matches are found for this piece, use “next” largest piece as the starting piece.

8. If the above steps have been completed (to finite exhaustion) and

(a) *there* **is** *one large remaining piece (i.e. the task* **is** *completed),* check that none of the pieces overlap eachother. If they fit and do not overlap, then done. If they do overlap, however, remove the pieces that do and attempt to swap them around (possibly more than two pieces). If they fit after swapping, then done. If not, then backup steps, ignoring pieces that previously fit in favor of others that are now “freed that may fit just as well. If so, continue in the forward direction once more.

(b) *there* **is not** *one large remaining piece (i.e. the task* **is not** *yet completed),* then repeat the process starting at step 2 but instead with non-straight-lined sides. Here the largest starting piece may be one individual piece or many that have already be matched together using the straight-line-edge search.

In presentation/interaction at the symposium, we will provide a description of where we are on these next two steps toward reaching Newell’s goal of integrated cognition via “One Program (= PERI) for Many Tasks.”

Acknowledgments

We’re indebted to ETS for sponsoring applied AI in the intersection of AI and testing (Phase 1 of the *e*WRITER system, bringing together Bringsjord, Phil Johnson-Laird, & Yingrui Yang), that led us to ponder the larger question of how to design and construct test-based integrated systems like PERI. In addition, we appreciate the willingness of Psychological Corporation to provide the sample WAIS test items shown in Figure 9.

References

Anderson, J. R., and Lebiere, C. 1998. *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum.

Anderson, J., and Lebiere, C. 2003. The newell test for a theory of cognition. *Behavioral and Brain Sciences* 26:587–640.

Anderson, J. R. 1993. *Rules of Mind*. Hillsdale, NJ: Lawrence Erlbaum.

Baron, R., and Kalsher, M. 2000. *Psychology*. Boston, MA: Allyn and Bacon.

Bringsjord, S., and Ferrucci, D. 1998. Logic and artificial intelligence: Divorced, still married, separated...? *Minds and Machines* 8:273–308.

Bringsjord, S., and Ferrucci, D. 2000. *Artificial Intelligence and Literary Creativity: Inside the Mind of Brutus, a Storytelling Machine*. Mahwah, NJ: Lawrence Erlbaum.

Bringsjord, S., and Schimanski, B. 2003. What is artificial intelligence? psychometric AI as an answer. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI–03)*, 887–893.

Bringsjord, S.; Arkoudas, K.; and Schimanski, B. forthcoming. Logic-based AI for the new millennium. *AI Magazine*.

Carpenter, P.; Just, M.; and Shell, P. 1990. What one intelligence test measures: A theoretical account of the processing in the Raven progressive matrices test. *Psychological Review* 97:404–431.

Cassimatis, N.; Trafton, J.; Schultz, A.; and Bugajska, M. 2004. Integrating cognition, perception and action through mental simulation in robots. In *Proceedings of the 2004 AAAI Spring Symposium on Knowledge Representation and Ontology for Autonomous Systems*.

Cassimatis, N. 2002. *Polyscheme: A Cognitive Architecture for Integrating Multiple Representation and Inference Schemes*. Ph.D. Dissertation, Massachusetts Institute of Technology (MIT).

Descartes, R. 1911. *The Philosophical Works of Descartes, Volume 1. Translated by Elizabeth S. Haldane and G.R.T. Ross*. Cambridge, UK: Cambridge University Press.

Ebbinghaus, H. D.; Flum, J.; and Thomas, W. 1984. *Mathematical Logic*. New York, NY: Springer-Verlag.

Evans, G. 1968. A program for the solution of a class of geometric-analogy intelligence-test questions. In Minsky, M., ed., *Semantic Information Processing*. Cambridge, MA: MIT Press. 271–353.

Fischler, M., and Firschein, O. 1987. *Intelligence: The Eye, the Brain, and the Computer*. Reading, MA: Addison-Wesley.

Genesereth, M., and Nilsson, N. 1987. *Logical Foundations of Artificial Intelligence*. Los Altos, CA: Morgan Kaufmann.

Ginsberg, M. 1993. *Essentials of Artificial Intelligence*. New York, NY: Morgan Kaufmann.

Harnad, S. 1991. Other bodies, other minds: A machine incarnation of an old philosophical problem. *Minds and Machines* 1(1):43–54.

Newell, A. 1973. You can’t play 20 questions with nature and win: Projective comments on the papers of this symposium. In Chase, W., ed., *Visual Information Processing*. New York: Academic Press. 283–308.

Nilsson, N. 1991. Logic and Artificial Intelligence. *Artificial Intelligence* 47:31–56.

Raven, J. C. 1962. *Advanced Progressive Matrices Set II*. London, UK: H. K. Lewis. Distributed in the United States by The Psychological Corporation Inc., San Antonio, Texas.

Rosenbloom, P.; Laird, J.; and Newell, A., eds. 1993. *The Soar Papers: Research on Integrated Intelligence*. Cambridge, MA: MIT Press.

Russell, S., and Norvig, P. 1994. *Artificial Intelligence: A Modern Approach*. Saddle River, NJ: Prentice Hall.

Russell, S., and Norvig, P. 2003. *Artificial Intelligence: A Modern Approach*. Upper Saddle River, NJ: Prentice Hall.

Spearman, C. 1927. *The Abilities of Man*. New York, NY: MacMillan.

Sternberg, R. 1988. *The Triarchic Mind: A New Theory of Human Intelligence*. New York, NY: Viking.

Stork, D. G. 1997. The end of an era, the beginning of another? HAL, Deep Blue, and Kasparov. http://www.research.ibm.com/deepblue/learn/html/e.8.1.html.

Sun, R. 2001. *Duality of the Mind*. Mahwah, NJ: Lawrence Erlbaum Associates.

Thurstone, L. 1938. *Primary Mental Abilities*. Chicago, IL: University of Chicago Press.

Torrance, E. P. 1988. The nature of creativity as manifest in its testing. In Sternberg, R., ed., *The Nature of Creativity*. Cambridge, UK: Cambridge University Press. 43–75.

Torrance, E. P. 1990. *The Torrance Tests of Creative Thinking*. Bensenville, IL: Scholastic Testing Service.

Turing, A. 1950. Computing machinery and intelligence. *Mind* LIX(236):433–460.

Wos, L.; Overbeek, R.; e. Lusk; and Boyle, J. 1992. *Automated Reasoning: Introduction and Applications*. New York, NY: McGraw Hill.