CREATIVE SOLUTIONS TO PROBLEMS

John McCarthy
Computer Science Department
Stanford University
Stanford, CA 94305
jmc@cs.stanford.edu
`http://www-formal.stanford.edu/jmc/`

March 29, 1999

Abstract

The idea is to chip a piece out of the problem of creativity by defining a *creative solution* to a problem relative to the functions and predicates used in posing the problem. The simplification comes from not talking about the creativity of the problem solver but only about the creativity of the solution.
Definition (informal): *A solution to a problem is creative if it involves concepts not present in statement of the problem and the general knowledge surrounding it.* Don't identify creativity with difficulty although they are usually correlated.
Example: The mutilated checkerboard problem.
We also consider how to express concisely the idea of a solution. Whether the expression is adequate is relative to the knowledge and ability of the person or program to which the idea is expressed.

1 Introduction

Making genuinely creative programs is likely to remain a distant goal for AI until someone comes up with a suitable new idea. Therefore, it is worthwhile to chip pieces off the creativity problem and work on them separately. It seems that it is possible to study the notion of a creative solution to a problem apart from studying how a person or machine finds the solution.

Definition (informal): *A solution to a problem is creative if it involves concepts not present in statement of the problem and the general knowledge surrounding it.*

As a small step toward programs that find creative solutions, we consider how to express the idea of a solution concisely. The adequacy of an idea of a solution is relative to the background of the person or program that will complete the solution. Conciseness isolates the idea from the background.

2 The mutilated checkerboard

Our first *Drosophila* for studying creative solutions is the *mutilated checkerboard problem.*

Two diagonally opposite corner squares are removed from a checkerboard. Is it possible to cover the remaining squares with dominoes? A domino is a 1×2 rectangle, i.e. can cover two rectilinearly adjacent squares.

The standard proof that this is impossible notes that a domino covers one black square and one white square, and therefore any covering by dominoes covers equal numbers of squares of the two colors. However, the mutilated checkerboard has 32 squares of one color and 30 squares of the other color.

We regard this proof as creative, because it involves an element not present in the formulation of the problem—namely the colors of the squares.

Here's a concise expression of the idea.

> A domino covers two squares of opposite color.

For some people this will be enough. Others may require the additional sentence

> The two squares that have been removed are of the same color.

One could argue that the colors are present implicitly when a checkerboard is mentioned, and perhaps the problem would be purer if it referred to an 8×8 array. However, one's initial reaction to the problem is to consider there being two colors as irrelevant, just as it is irrelevant what the two colors are—some boards have black and white squares, some have red and black, and some have green and off-white (perhaps thought to be more relaxing for the players).

(McCarthy 1964) presented the problem as a tough nut for first order theorem provers, because a straightforward formalization of the problem provides no way of making the argument in the language of the formalization. Creativity was not mentioned in that 1964 memo, but I evidently mentioned it in lectures, because the problem gave rise to an informal competition aimed at finding the most non-creative solution to the problem.

The first "non-creative" solution was proposed by Shmuel Winograd of IBM. He claimed it was non-creative, because it didn't involve coloring.

Assume a covering. The number of dominoes projecting from the top row to the second row is odd. Likewise the number from the second to the third is odd, etc. Therefore, the total number of vertical dominoes is the sum of seven odd numbers and hence odd. Likewise the the number of horizontal dominoes is odd. Odd + odd is even so the total is even, but the total is 31. There is an apparent mathematical induction here in the "etc.". We will see later that the idea itself does not include the induction.

Surely Winograd's proof is creative, but we can ask whether there is one creative idea in it or several. My guess is that there was one creative idea, and the rest was straightforward for a good mathematician like Winograd.

The second was by Marvin Minsky of M.I.T.

Start with the 2-diagonal next to an excluded corner square. 2 dominoes must project from it to the adjacent 3-diagonal. Subtracting, one domino projects from the 3-diagonal to the 4-diagonal. 3 project from the 4-diagonal to the 5-diagonal, 2 project from the 5-diagonal to the 6-diagonal, 4 from the 6-diagonal to the 7-diagonal and finally 3 project from the 7 diagonal to the 8-diagonal. This covers only 3 of the 8 squares in the 8-diagonal. Coming from the opposite excluded corner also only covers 3 squares of the 8-diagonal, leaving 2 uncovered squares. Minsky's proof gets high points for non-creativity, because it is specific to the 8 by 8 board.

The third method is by Dimitri Stefanuk of Moscow, Russia. He suggested 62 proofs—or 17, taking into account symmetries.

Choose an arbitrary square and mark it 1. Mark its rectilinear neighbors 2, their unmarked neighbors 3, continuing until every unexcluded square is marked. Then proceed as in Minsky's proof, counting the number of dominoes projecting from the 1-squares to the 2-squares, etc. We get a proof if there are not enough dominoes projecting from the $n-1$ squares to the n-squares. Every attempt at a Stefanuk proof succeeds. Stefanuk proofs are just as uncreative a Minsky proofs.

Remark: Counting colors shows that Stefanuk proofs and Minsky proofs always work on any even board. However, this argument is at a meta-level to the Minsky and Stefanuk proofs and therefore can't claim to be non-creative.

In fact, the Winograd, Minsky and Stefanuk proofs are all creative, and we will try to identify the creativity involved and give a concise expression of the ideas.

The most recent proofs, by separate computer programs written by Mark Stickel ((Uribe and Stickel 1994)) and Dan Pehoushek (personal communication) apply propositional satisfaction algorithms to a propositional expression in 256 variables. There is a variable asserting for each square and each of the four direction whether a domino projects from that square in that direction.[1]

The propositional satisfaction proofs are genuinely non-creative, and but because they come from computer programs. Writing the propositional calculus formulas asserting that a domino projects in exactly one direction from each square, that there is a domino project back from a square into which a domino projects and that no domino projects off the board or into a forbidden square is taken as straightforward, i.e. not involving creativity.

(Subramanian 1993) includes an interactive proof using the Boyer-Moore prover NQTHM. Allen Newell also discussed the problem.

3 Pinning down the ideas, English first

Actually solving a problem involves both a creative part and a routine part. It is sometimes possible to separate the creative part from the routine part. Consider the Minsky solution. We can tell a person,

> Starting with the diagonal next to an excluded square, compute how many dominoes project from each diagonal to the next diagonal.

[1]Stickel's program solved it in 80.3 seconds in 1983 and running on a Pentium 200 solved it recently in 5.3 seconds.

This proved to be adequate for a sample of one veterinarian. This is a kind of program, but it doesn't need a termination condition, because the termination is apparent to the problem solver. It is harder to tell whether it is close to the form in which the solution was discovered, although it looks close.

When we consider expressing the creative idea to a computer program, it would be most convenient to tell it a fact or to tell it an executable program. I know of no-one who has tried to make computer programs that can accept a fragment of a program as above as the non-routine part of the solution to the problem.

Here's a try at expressing concisely the idea of the Winograd solution. It seems to be longer and to involve more than one idea.

> Starting with the top row, compute the parity of the number of dominoes projecting down out of each row. Consider the parity of the sum. Repeat going horizontally starting with the left column. Compute the parity of the total number of dominoes compared to the sum of the two parities.

According to a 1999 personal communication, the idea can be expressed more concisely. Winograd got the idea as a result of seeing a movie about the Socratic method in mathematics education. A student said something like, "An odd number of dominoes project down from the top row and an odd number down from the second row to the third . . . ". At this point the teacher interrupted the student and led him by the Socratic method to the standard solution. Winograd wondered whether the student's original idea could be pushed through, and this led to the solution.

Looking at it this way, the creative idea seems to be

> Note that an odd number of dominoes project from the top row to the second and continue from there.

Presumably not everyone would be able to push the argument through, because at some point you have to notice that although you get no contradiction from determining that there are an odd number of vertical dominoes, you also get an odd number of horizontal dominoes and an even number of dominoes altogether. Thus the notion of *creative solution* seems to depend on the ability available to push an idea through.

Winograd idea can also be pushed through when two squares of the same color are removed any where on the board. If one starts at the top and computes the successive parities, there is a jog in parity when an excluded square is encountered. Taking into account the jog one gets an answer as to the parity of the number of vertical dominoes. For some pairs of excluded squares, the parity turns out to be even. However, the parity turns out, it will be the same going horizontally.

The Winograd idea can then be expressed more abstractly.

> Successively compute the parity of the number of vertical dominoes projecting from each row to the next.

The Stefanuk solution seems to involve two ideas. The first is to label the squares starting from an arbitary square. The second is like the Minsky solution, namely

> Starting with $n = 1$, compute how many dominoes project from the set of squares labelled n to the squares labelled $n + 1$.

In this section I have striven for concise expressions of the creative idea. The reason is to establish that there is one idea in each of the cases. Obviously it will be easier to make computers come up with creative solutions if the idea of the solution is just one thing—whatever kind of thing that may be.

4 Elementary first order formulations

The impossibility statement is readily formulated as a sentence of the predicate calculus, but I don't see how the parity and counting argument can be translated into a guide to the method of semantic tableaus, into a resolution argument, or into a standard proof.

The word "elementary" is used in the sense that the quantifiers range over numbers and not over sets.

The formulas are here for completeness. If you don't like reading them, the rest of this section may be skipped.

We number the rows and columns from 1 to 8 and we introduce predicates $S(x, y)$, $L(x, y)$, $E(x, y)$, $G^1(x, y)$, $G^2(x, y)$, $G^3(x, y)$, $G^4(x, y)$, and $G^5(x, y)$ with the following intended interpretations: $S(x, y)$ means $y = x + 1$.

$L(x, y)$ means $x < y$.
$E(x, y)$ means $x = y$.

$G^1(x,y)$ means the square (x,y) and the square $(x+1,y)$ are covered by a domino.

$G^2(x,y)$ means the square (x,y) and the square (x,y+1) are covered by a domino.

$G^3(x,y)$ means the square (x,y) and the square (x-1,y) are covered by a domino.

$G^4(x,y)$ means the square (x,y) and the square (x,y-1) are covered by a domino.

$G^5(x,y)$ means the square (x,y) is not covered. We shall axiomatize only as much of the properties of the numbers from 1 to 8 as we shall need.

1. $S(1,2) \wedge S(2,3) \wedge S(3,4) \wedge S(4,5) \wedge S(5,6) \wedge S(6,7) \wedge S(7,8)$
2. $S(x,y) \rightarrow L(x,y)$.
3. $L(x,y) \wedge L(y,z) \rightarrow L(x,z) \wedge \neg S(x,z)$.
4. $L(x,y) \rightarrow \neg E(x,y)$.
5. $E(x,x)$.

   These axioms insure that all eight numbers are different and determine the values of $S(x,y)$, $L(x,y)$, and $E(x,y)$ for $x,y = 1,\ldots,8$.
6. $G^1(x,y) \vee G^2(x,y) \vee G^3(x,y) \vee G^4(x,y) \vee G^5(x,y)$
7. $G^1(x,y) \rightarrow \neg(G^2(x,y) \vee G^3(x,y) \vee G^4(x,y) \vee G^5(x,y))$
8. $G^2(x,y) \rightarrow \neg(G^3(x,y) \vee G^4(x,y) \vee G^5(x,y))$
9. $G^3(x,y) \rightarrow \neg(G^4(x,y) \vee G^5(x,y))$
10. $G^4(x,y) \rightarrow \neg G^5(x,y)$

    These axioms insure that every square (x,y) satisfies exactly one $G^i(x,y)$.
11. $G^5(x,y) \equiv x = 1 \wedge y = 1 \vee x = 8 \wedge y = 8$.
12. $G^5(x,y) \rightarrow (E(1,x) \wedge E(1,y)) \vee (E(8,x) \wedge E(8,y))$ These axioms insure that the uncovered squares are precisely (1,1) and (8,8).
13. $S(x_1,x_2) \rightarrow G(x_1,y) \equiv G^3(x_2,y)$
14. $S(y_1,y_2) \rightarrow G^2(x,y_1) \equiv G^4(x,y_2)$ These axioms state the conditions that a pair of adjacent squares be covered by a domino.
15. $\neg G^3(1,y) \wedge \neg G^1(8,y) \wedge \neg G^2(x.8) \wedge \neg G^4(x,1)$

These axioms state that the dominoes don't stick out over the edge of the board.

Suppose we had a model of these 15 sentences (in Robinson's clausal formalism, there would be 31 clauses). There would have to be eight individuals $1', \ldots, 8'$ satisfying the relations asserted for $1, \ldots, 8$ in the axioms. They would have to be distinct since axioms 1,2, and 3 allow us to prove $L(x, y)$ whenever this is so and axioms 4 and 5 then allow us to show that $L(x, y)$ holds only for distinct x and y.

We then label the squares of a checkboard and place a domino on each square (x,y) that satisfies $G^1(x, y)$ or $G^2(x, y)$ sticking to the right or up as the case may be. Axioms 13 and 14 insure that the dominoes don't overlap, axioms 6-12 insure that all squares but the corner squares are covered and axiom 15 insures that no dominoes stick out over the edge.

Since there is no such covering the sentences have no model and are inconsistent.

In a formalism that allows functions and equality we have a briefer inconsistent set of sentences involving

$s(x)$ the successor of x

$g(x, y)$ has value of 1 to 5 according to whether $G^1(x, y)$ or ... or $G^5(x, y)$

The sentences are

1. $s(s(s(s(s(s(s(s(8)))))))) = 8$.
2. $\neg s(s(s(s(x)))) = x$ The sentences insure the existence of 8 distinct individuals using a cyclic successor function.
3. $g(x, y) = 5 \equiv x = 8 \wedge y = 8 \vee x = 1 \wedge y = 1$ Insures that exactly the corner squares (1,1) and (8,8) are uncovered.
4. $g(x, y) = 1 \equiv g(s(x), y) = 3$
5. $g(x, y) = 2 \equiv g(x, s(y)) = 4$ Each domino covers two adjacent squares
6. $g(1, y) \neq 3 \wedge g(8, y) \neq 1 \wedge g(x, 1) \neq 4 \wedge g(x, 8) \neq 2$
   Dominoes don't stick out
7. $1 = s(8) \wedge 2 = s(1) \wedge 3 = s(2) \wedge 4 = s(3) \wedge 5 = s(4)$.
8. $g(x, y) = 1 \vee g(x, y) = 2 \vee g(x, y) = 3 \vee g(x, y) = 4 \vee g(x, y) = 5$.

This identifies the numbers used and ties down the values of $g(x, y)$.

Not only is this language incapable of expressing the colors of the squares. It is also incapable of expressing the counts of the numbers of squares with a given property, the latter being required for expressing the Minsky, Winograd and Stefanuk proofs.

5 Expressing the creative ideas in set theory

In this section we try to identify the creative aspects of the proofs with specific formulas of logic. We use Zermelo-Fraenkel set theory, formalized in first order logic, because its power allows formulas that are closer to expressing the human ideas behind the proofs.

The traditional proof and the proofs by Winograd, Minsky and Stefanuk can all be based on the following sentences of set theory axiomatized in first order logic.

5.1 Definitions common to the proofs

We have the definitions

$$Board = Z8 \times Z8, \tag{1}$$

$$mutilated\text{-}board = Board - \{(0,0),(7,7)\}, \tag{2}$$

$$\begin{aligned}&domino\text{-}on\text{-}board(x) \equiv (x \subset Board) \wedge card(x) = 2\\ &\wedge(\forall x1\ x2)(x1 \neq x2 \wedge x1 \in x \wedge x2 \in x\\ &\rightarrow adjacent(x1, x2))\end{aligned} \tag{3}$$

and

$$\begin{aligned}&adjacent(x1, x2) \equiv |c(x1,1) - c(x2,1)| = 1\\ &\wedge c(x1,2) = c(x2,2)\\ &\vee |c(x1,2) - c(x2,2)| = 1 \wedge c(x1,1) = c(x2,1).\end{aligned} \tag{4}$$

If we are willing to be slightly tricky, we can write more compactly

$$\begin{aligned}&adjacent(x1, x2) \equiv\\ &|c(x1,1) - c(x2,1)| + |c(x1,2) - c(x2,2)| = 1,\end{aligned} \tag{5}$$

but then the proof might be more difficult for a computer program.

Next we have.

$$\begin{aligned}&partial\text{-}covering(z)\\&\equiv (\forall x)(x \in z \rightarrow domino\text{-}on\text{-}board(x))\\&\wedge(\forall x\ y)(x \in z \wedge y \in z \rightarrow x = y \vee x \cap y = \{\})\end{aligned} \tag{6}$$

Theorem:

$$\neg(\exists z)(partial\text{-}covering(z) \wedge \bigcup z = mutilated\text{-}board) \tag{7}$$

Up to this point, there is no creativity. Someone might argue that the decision to use set theory is creative, but I'm striving for a technical notion of *creative solution*, and one has to start with some background language. Taking set theory as background is more likely to lead to technical results than trying to use English.

The proofs are as follows:

5.2 Standard proof

We define

$$x \in Board \rightarrow color(x) = rem(c(x,1) + c(x,2), 2). \tag{8}$$

Making this definition is a creative step, but it doesn't give the solution by itself.

$$\begin{aligned}&domino\text{-}on\text{-}board(x) \rightarrow\\&(\exists u\ v)(u \in x \wedge v \in x \wedge color(u) = 0\\&\wedge color(v) = 1).\end{aligned} \tag{9}$$

$$\begin{aligned}&partial\text{-}covering(z) \rightarrow\\&card(\{u \in \bigcup z | color(u) = 0\})\\&= card(\{u \in \bigcup z | color(u) = 1\}).\end{aligned} \tag{10}$$

This is the key creative step.

$$\begin{aligned}&card(\{u \in mutilated\text{-}board | color(u) = 0\})\\&\neq card(\{u \in mutilated\text{-}board | color(u) = 1\}),\end{aligned} \tag{11}$$

and finally

$$\begin{aligned}&\neg(\exists z)(partial\text{-}covering(z)\\&\wedge mutilated\text{-}board = \bigcup z)\end{aligned} \tag{12}$$

Q.E.D.

This last step is presumably routine.

(McCarthy 1996) argued that the above proof, or some equally concise expression of its ideas, should be accepted by any mathematical proof checker that mathematicians would actually use. Unfortunately, no present proof checker comes close.

The following two definitions are used in the proofs by Winograd, Minsky and Stefanuk.

$$\begin{aligned}&(\forall x\ A\ B)(\mathit{domino\text{-}on\text{-}board}(x)\\&\wedge A \subset \mathit{Board} \wedge B \subset \mathit{Board} \rightarrow\\&(\mathit{overlaps}(x, A, B) \equiv\\&x \cap A \neq \{\} \wedge x \cap B \neq \{\})\end{aligned} \tag{13}$$

and

$$\begin{aligned}&(\forall A\ B)(A \subset \mathit{Board} \wedge B \subset \mathit{Board}\\&\rightarrow (\mathit{size\text{-}overlap}(A, B) =\\&\mathit{card}(\{x|\mathit{overlaps}(x, A, B)\}))).\end{aligned} \tag{14}$$

I suppose they count as creative, but maybe as one creation rather than two.

In a future article, we hope to put the Winograd and Stefanuk proofs in a logical form that isolates the creative part from the routine part.

Acknowledgements

I am indebted to Tom Costello for both scientific and TeXnical help.

The work was supported by the Air Force Office of Scientific Research New World Vistas program and by the Defense Advanced Research Projects Agency High Performance Knowledge Bases program.

References

McCarthy, J. 1964. **a tough nut for theorem provers**[2]. Stanford AI Memo 16—now on the web.

McCarthy, J. 1996. **the mutilated checkerboard in set theory**[3]. presented at a 1996 conference in Warsaw.

[2]http://www-formal.stanford.edu/jmc/nut.html
[3]http://www-formal.stanford.edu/jmc/checkerboard.html

Subramanian, S. 1993. *A Mechanized Framework for Specifying Problem Domains and Verifying Plans.* PhD thesis, University of Texas, Austin.

Uribe, T., and M. Stickel. 1994. Ordered binary decision diagrams and the davis-putnam procedure. In J. P. Drummond (Ed.), *Lecture Notes on Computer Science, First international conference on constraints in computational logics*, Vol. 845. Springer, September.

/@steam.stanford.edu:/u/jmc/w99/creative.tex: begun Sat Mar 6 20:57:30 1999, latexed March 29, 1999 at 3:18 p.m.