

The Marriage Problem: How to Choose?

Bruce Brown¹

1. Introduction

The problem of how to successfully choose the partner most likely to lead to a long and happy marriage is a task which has occupied the minds of young and older people alike, men and women, among all races and cultures, throughout the ages. Even though modern social conditions in our society have created a lot of flexibility in the entering and leaving of family partnership arrangements, many of the more traditional attitudes towards marriage still remain as valid as ever. So, there could well be something worth thinking about in studying the solution to the so-called *marriage problem*, a very old problem in Probability Theory.

Indeed, the solution to this problem is of interest from several viewpoints, not just for the rather simple suggestion it makes about choosing a marriage partner. The method of solution uses elements from the theory of optimal stopping, a sequential decision-making process which finds application in many areas, the most obvious being in medical clinical trials with strong connections to dynamic programming, and other modern optimisation methods. The solution to the marriage problem is straightforward as the reader will soon see but indicates also how extensions can be made to more complex situations, and what the difficulties are of so doing.

The problem has several equivalent forms, going under different names, some of which, like the *secretary problem*, are a little outdated. The task is easily stated: faced with a choice among n candidates, who present themselves for evaluation in random order, how to choose the best among them? Or, in probabilistic terms, what decision rule should be used to maximise the probability of choosing the best candidate?

The instinctive reaction to this question is just to say, evaluate all of them and then choose the best one. But of course, courting doesn't work that way. In this situation, there are severe constraints on the decision making process. It is necessary to come to a decision, whether to choose or reject each candidate, immediately after the individual evaluation has been performed. And once made, that decision to accept or reject is final and absolute. Acceptance is acceptance for life, with no changing of mind allowed, and rejection is rejection without any possibility of later recall.

In the world of the marriage problem, which we denote by WMP, you cannot get married and then casually, a short time later, tell your new husband or wife that you've found someone better and intend to move on, even though in real life some people do behave exactly this way. And you can't reject someone, let's say Sally, then proceed to take the time to 'evaluate' a whole string of subsequent partners, all of whom turn

¹Professor Bruce Brown is a professor in statistics in the School of Mathematics and Statistics at the University of New South Wales.

out not to be as good as Sally, then later, upon realising this, give Sally a call and offer to reverse the original rejection. In WMP this can't be done, and Sally's reply isn't printable even though, once again, you can see that some people do things this way in real life.

So the reader may object, and say that in real life prospective partners change their minds all the time, and that forgiveness and acceptance often happen, or that de facto arrangements, divorce and re- marriages all make the real picture much more complicated. And all of those objections are no doubt true: the marriage problem as stated so far seems to derive from an earlier time, when social rules were simpler and more rigid than now.

Not only are divorces and re-marriages implicitly non-existent in WMP, but there is another unrealistic aspect: it implies an attitude to life which is too egotistical. To see this, consider the goal: to maximise the probability of choosing the best candidate. This maximises the expected value of a very simple underlying reward structure; one unit for having the satisfaction, at the end of one's long life, of realising that your chosen spouse was the very best person among all the other possibilities among the opposite sex that you ever encountered; and zero if this is not the case.

So, even though your married life may have been long and happy, the fact that, among all the prospective life partners you encountered along the way, the prospect that one or two others might have been 'better' in some way is enough to drive you to the despair of according yourself a reward of zero for your whole life's married experience! How absurd is this? It is a statement of vanity out of control in an unrealistic way.

Yet, despite all these reservations about the unworldly WMP assumptions behind the marriage problem, it may be that still, the mathematical solution gives us something useful to think about.

What follows next is an informal proof of the solution. The reader interested only in the solution itself and its implications will have to skip to Section 3.

2. Solving the Marriage Problem

The problem can be presented from either the male or female point of view. Also, there are two cases to consider, small n and large n . So let us phrase the small n solution from the male viewpoint and the large n solution from the female viewpoint.

The small n case

We are considering the situation of a young man who thinks that he will not meet many young women who would agree to marry him. If he thinks that $n = 1$, that he will only ever meet one girl he could marry, then the decision process is simple - when that happens, propose to her without delay. But if he thinks that $n = 2$, then when the first such prospect comes along, it is a completely random choice whether to accept her, or wait for the second prospect, who has a 50:50 chance of being better than the first. From a purely practical viewpoint, if there is some doubt about whether the second prospect would ever materialise, then the best policy would be to propose to the first girl.

The maximised probability of choosing the best future wife is 1 when $n = 1$, and

$\frac{1}{2}$ when $n = 2$. In both cases the probability has the form $\frac{1}{n}$ which is the success probability from making a purely random, unintelligent choice. It is only when we get to the case $n = 3$ that things begin to get more interesting.

What are the possible decision rules when $n = 3$? Keep in mind that there is never any point in accepting a prospective wife if she is worse than one previously rejected, because then the probability of choosing the best is known to be zero. The possible rules are as follows.

(i) Propose to the first girl.

(ii) Reject the first girl no matter how attractive she may be and then propose to the next girl if she is better than the first, rejected prospect. But if she is not better than the first one, then it is necessary to proceed to the third, and hope that she will prove to be the best of all.

(iii) Reject the first two girls no matter how appealing they may be and propose to the third.

Now there are six possible orderings for the sequence of rankings of the three girls, as they present themselves for evaluation. It is reasonably easy to write these down, and then to calculate the probability of choosing the best prospect among the three, using decision rules (i), (ii) or (iii) above. The probabilities are $\frac{1}{3}, \frac{1}{2}, \frac{1}{3}$ respectively. Thus decision rule (ii) is the best, and has a success probability of $\frac{1}{2}$, better than the value $\frac{1}{n} = \frac{1}{3}$ for a random choice.

The case $n = 4$ can be analysed in a similar manner. Now there are more possible decision rules, as follows.

(i) Propose to the first girl.

(ii) Reject the first girl, and propose to the second if she is better than the first. If not, proceed to the third, and propose to her if she is better than the first two. But if this also is not the case, it is necessary to proceed to the last, the fourth prospect in the hope that she proves to be the best one overall.

(iii) Reject both of the first two prospects no matter how appealing they may be. Then accept number three if she is better than the first two, or otherwise, pass on in hope to the fourth.

(iv) Reject the first three, and choose the fourth.

The success probabilities of these four strategies are $\frac{1}{4}, \frac{11}{24}, \frac{10}{24}, \frac{1}{4}$ respectively. So the best decision rule is to reject the first girl, and propose to the next one who is better than all previous candidates. This success probability $\frac{11}{24}$ is clearly greater than the value $\frac{1}{n} = \frac{1}{4}$ for a random choice.

As we examine higher values of n , it becomes easier to use an algebraic argument. This is outlined in the next section where the case of large n is examined. For the case $n = 5$, the success probabilities for rejecting either 0, 1, 2, 3 or 4 girls before then proposing to the next who is better than all previous, are, respectively, $\frac{1}{5}, \frac{5}{12}, \frac{13}{30}, \frac{7}{20}, \frac{1}{5}$.

So the best strategy is to reject the first two prospects, and then propose to the next one who is better than all those previously examined. Once again the success probability $\frac{13}{30}$ is clearly greater than the value $\frac{1}{n} = \frac{1}{5}$ for a random choice.

The case of large n .

Let us discuss this case from the female viewpoint, of a young woman who knows that n proposals of marriage would come her way during the allotted time she has chosen, at the end of which period she definitely wants to be married. We will assume that the potential value of n is reasonably large, something that applies to many young women, as long as the allotted time period for a decision is not too short.

The task is to find the optimal strategy to maximise the probability of choosing the best among n suitors whose marriage proposals come along in a random order-of-merit. At each stage, the only possible actions are to accept, a lifetime decision, or to reject without any chance of a later recall. The case where n is large has a nice solution, outlined as follows.

(i) Reward structure. Notice that maximising the probability of choosing the best is equivalent to maximising the expected reward, where the reward structure is: 1 for choosing the best, and zero otherwise.

(ii) Suppose that the first $j - 1$ suitors have been rejected, and the j th offer is being considered. The question is whether to stop now (i.e. accept) or continue (i.e. reject). Intuitively, we can see that the best action is to stop if our expected reward for stopping now exceeds v_j , the maximal expected reward for continuing beyond the j th suitor.

Our expected reward for stopping now is zero if the present suitor is not the best so far, because then we know that a better suitor exists. But if the present suitor is best so far, the probability is $\frac{j}{n}$ that he is best overall, i.e. that the best one occurs in the first set of j suitors rather than the last $n - j$ suitors. Therefore, the expected reward for stopping now is: 0 if the j th suitor is not best so far; $\frac{j}{n}$ if the j th suitor is best so far.

Note that $\frac{j}{n}$ is an increasing function of j .

(iii) This has to be compared with v_j , the maximal expected reward for continuing, but the form of v_j is unknown. However, we can observe that v_j must be a *decreasing* function of j , because the set of strategies for continuing beyond the j th suitor includes all the strategies for continuing beyond j' , when $j < j'$.

So, in comparing the reward for stopping now with the maximal expected reward for continuing, we are comparing an *increasing* function with a *decreasing* function. Clearly, two such functions will cross over at one, and only one place. Thus, there will be just one integer value j^* such that $\frac{j}{n} > v_j$ for all $j > j^*$. Therefore the optimal stopping rule is:

Reject all suitors for $j \leq j^*$. Accept the first suitor for $j > j^*$ who is the best so far.

(iv) The only difficulty is that the value of the threshold j^* is unknown. However, we can do the calculations for a general value of j^* , and choose the actual value to maximise the probability of choosing the best suitor. This exercise requires some serious combinatorial probability calculations.

(v) Write $j^* = k$ for notational convenience. The probability of choosing the best suitor is the probability that

(a) the best suitor is not among the first k rejected suitors, and (b) if the best among the first k suitors has rank $r \geq 2$ among all the suitors, then the suitor with rank 1 comes along before any of the other $r - 2$ suitors who are all better than the best among the first k .

The probability in (b) is $\left(\frac{1}{r-1}\right)$. This has to be summed over all the possible values of $r \geq 2$, after being weighted by the probabilities for those r values.

(vi) The probability of a particular r value is the probability that the best $r - 1$ all get allocated to the last batch of $n - k$ suitors, and that the r th suitor gets allocated to the first batch of size k . Using the chain rule of probability (just as in the calculation of one player getting four aces in a hand of bridge as $\frac{13}{52} \times \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49}$), this probability is

$$\frac{(n-k)(n-k-1)(n-k-2)\dots(n-k-r+2)}{n(n-1)(n-2)\dots(n-r+2)} \times \frac{k}{(n-r+1)}.$$

The required probability now is the sum

$$\sum_{r=2}^{n-k+1} \frac{(n-k)(n-k-1)(n-k-2)\dots(n-k-r+2)}{n(n-1)(n-2)\dots(n-r+2)} \times \frac{k}{(n-r+1)} \times \left[\frac{1}{r-1}\right].$$

This looks like a big and complicated sum, but it has a nice approximation if n is large, and $x = k/n$ is the fraction of automatically rejected suitors. Then it is approximately $\sum_{r=2}^{\infty} \frac{x^r}{r-1}$, which is $x - \log x$. It is a simple exercise in calculus to show that this is maximised by the choice $x = e^{-1} = 0.368$. Thus, in the optimal rule, the first 36.8% of suitors are rejected, and the next one who is best so far is accepted. And the optimal success probability is also $e^{-1} = 0.368$, considerably larger than $\frac{1}{n}$ for a random choice, when n is large.

But, even using the optimal strategy, there is still a probability of around 63% of failing to select the best partner.

3. Applying the rule in practice

A frequently asked question is: how can I know what is the value of n that applies to me. Probably the most rational way to deal with this is to assume that n is large, and then, rather than worry about what its actual value might be, allocate a time period, at the end of which you want the marriage question to be resolved.

To illustrate, consider a young man who is willing to consider young women as marriage prospects once he reaches the age of 20, and wants to have the marriage question settled before he turns 39, a time period of 19 years. Now 36.8% of 19 is very nearly 7, so his optimal strategy is:

- (i) do not propose to any girl, no matter how enticing, before reaching the age of 27;
- (ii) after turning 27, propose to the very next girl who is better than all previous prospects.

Despite the reservations about the unworldly assumptions behind the marriage problem, you can see that this is quite a reasonable approach, and that this, or something similar, is likely to be followed by many young men. The same applies to the corresponding rule for women, where the allotted individual time periods may be earlier, and shorter, than the ones chosen by men.

It is quite interesting to know that if you question older people carefully, you often find that they have used something quite close to the marriage problem solution in choosing their spouse. But take care! You have to conduct such conversations with some delicacy, because not everyone wants to admit the existence of past loves, mistakenly rejected, whose memory may still burn brightly.

As for the author, he can tell you that, looking back and doing some calculations, he did follow the marriage problem solution, albeit by accident, and that it has all worked out perfectly.