

Programming Abstractions

Week 6: Modules, data types, and backtracking

Stephen Checkoway

Modules in Racket

(Lack of) modules in Scheme

Traditional

Scheme has a `(load file-name)` that's like C's `#include`

- ▶ `(load "foo.scm")` simply reads in the content of the file `foo.scm` as scheme code

Upsides

- ▶ It's simple: It simply makes all of the definitions in the loaded file available in the current file

Downsides

- ▶ Only a single namespace so different files can't have procedures with the same name
- ▶ Allows no separation between an interface and an implementation
 - E.g., "private" helper functions aren't actually private

Modules in Racket

Modern

Each file that starts with `#lang` creates a module named after the file

`#lang` also specifies the language of the file

- ▶ Racket was designed to implement programming languages
- ▶ We're only going to use the Racket language itself
- ▶ All of our files start with `#lang racket`

Exposing definitions

`(provide ...)`

By default, each definition you make in a Racket file is private to the file

To expose the definition, you use `(provide ...)`

To expose all definitions, you use
`(provide (all-defined-out))`

E.g.,

```
#lang racket
(provide (all-defined-out))
```

```
(define (mul2 x)
  (* x 2))
```

This explains the line in the homework

```
; Export all of our top-level definitions so that tests.rkt  
; can import them. See tests.rkt.  
(provide (all-defined-out))
```

Exposing only some definitions

`(provide sym1 sym2...)`

You can specify exactly which definitions are exposed by specifying them via one or more `provides`

```
#lang racket
```

```
(provide foo-a foo-b)
```

```
(provide bar-a bar-b)
```

```
(define helper ...) ; Not exposed
```

```
(define foo-a ...)
```

```
(define foo-b ...)
```

```
(define bar-a ...)
```

```
(define bar-b ...)
```

Importing definitions from modules

`(require ...)`

To get access to a module's definitions we need to `require` the module

E.g., the `test.rkt` files in the assignments require the homework file

▸ `(require "hw1.rkt")` imports the definitions from the file `hw1.rkt`

Imagine you're writing code to analyze DNA sequence data so you create a Racket file `analysis.rkt`. You write two procedures `dna-match` and `edit-distance`. The `edit-distance` procedure is just a helper procedure and is only used inside of `dna-match`. What code should you add to `analysis.rkt` to expose the appropriate procedures?

A. `(provide (all-defined-out))`

B. `(provide dna-match)`

C. `(provide edit-distance)`

D. `(hide edit-distance)`

E. `(provide dna-match)`
`(hide edit-distance)`

Now, you're writing another module (i.e., another file) that wants to use the `dna-match` procedure in the `analysis.rkt` file. What code should you use to make that procedure available for use inside the current module?

A. `(import dna-match)`

B. `(require dna-match)`

C. `(import "analysis.rkt")`

D. `(require "analysis.rkt")`

E. `(import-all "analysis.rkt")`

Plus a whole lot more

Racket supports a dizzying array of module options

`#lang racket` is a shorthand for `(module module-id racket ...)`

Submodules can be created using a variety of module forms: `module`, `module*`, `module+`

Requiring modules can

- ▶ Import specific symbols
- ▶ Exclude specific symbols
- ▶ Rename symbols (e.g., two modules can be imported with conflicting names)

We won't need any of this extra functionality in the course, because our programs are so short

Data types

Data types

We're going to construct data types out of lists (of course)

The first element in the list is going to be a symbol that's the name of the data type

Depending on the specific data type, the other elements in the list will either be the elements that comprise an instance or fields of the data type

What do we need to implement a data type?

A representation for the data type: a list with a particular structure

Procedures to work with an instance of the data type

- ▶ Recognizers: Is this thing an object of type X?
- ▶ Constructors: Create an object of type X
- ▶ Accessors: Get field Y from an object of type X

Special values (not applicable to all data types)

Since we're working functionally, we don't need to *set* any fields, we would just create a new object with the appropriate field

Representation

We're going to use lists to represent instances of a data type

Example: A `set` data type which will hold a (mathematical) set of values for us

Empty set: let's use the list `'(set)` (equivalently `(list 'set)`)

Nonempty set: let's take the list of elements (with no duplicates) and cons `'set` on the front

- ▶ `'(set 1 3 5 7 9)`
- ▶ `'(set a)`
- ▶ `'(set x z y)`

Using the name of the data type as the first element of the list is traditional

Recognizers

Recognizers are procedures that return `#t` or `#f` corresponding to whether or not the passed in object is of the appropriate type

- ▶ These are analogous to `number?` and `list?`

There are also recognizers that return `#t` or `#f` corresponding to whether or not the passed in object has a particular value of the type

- ▶ These are analogous to `zero?` and `empty?`

Recognizers for our set data type

We want to know if a particular object is a set so we'll write a procedure `set?`

```
(define (set? obj)
  (and (list? obj)
       (eq? (first obj) 'set)))
```

This is analogous to `list?` except it returns `#t` if the object is a set

Just as `(empty? x)` returns `#t` if `x` is an empty list, let's write `(empty-set? x)` which returns `#t` if `x` is an empty set.

Remember, we're representing a set as a list starting with `'set` where the rest of the elements of the list are elements of the set. How do we do this?

A.

```
(define (empty-set? obj)
  (= (length obj) 1))
```

D. Any of A, B, or C

E. Either B or C

B.

```
(define (empty-set? obj)
  (and (= (first obj) 'set)
        (empty? (rest obj))))
```

C.

```
(define (empty-set? obj)
  (and (set? obj)
        (empty? (rest obj))))
```

Constructors

Now that we know how to recognize if something is an instance of our data type, we need procedures to create them

Typically, we use the name of the data type itself

Example:

- ▶ To create a set, we need a list of elements
- ▶ The list might have duplicates, so we should remove those

```
(define (set elements)
  (cons 'set (remove-duplicates elements)))
```

Special value for our set data type

Just as list has a special value, empty, it might be nice to have an empty-set

```
(define empty-set (set empty))
```

Accessors

We need a way to access the sub-objects of an instance of our data type

If we think about Python or Java objects, they have fields. We can do something similar in Racket (and will in just a few slides) and we'll need procedures to access them

For our set example, we have a list of elements and we would like to return that list

Set accessor

```
(define (set-members s)
  (if (set? s)
      (rest s)
      (error 'set-members "~v is not a set" s)))
```

There are multiple forms of the (error ...) procedure, this one is
(error procedure-name format-string arguments)

The ~v means to substitute a string representation of the object for the ~v

```
> (set-members '(1 2 3))
set-members: '(1 2 3) is not a set
```

Example: set

```
(define (set elements)
  (cons 'set (remove-duplicates elements)))

(define (set? obj)
  (and (list? obj)
       (eq? (first obj) 'set)))

(define (empty-set? obj)
  (and (set? obj)
       (empty? (rest obj))))

(define (set-members s)
  (if (set? s)
      (rest s)
      (error 'set-members "~v is not a set" s)))

(define empty-set (set empty))
```

Additional procedures

```
(define (set-contains? x s)
  (member x (set-members s)))
```

```
(define (set-insert x s)
  (if (set-contains? x s)
      s
      (cons 'set (cons x (set-members s)))))
```

; We could have used (set (cons x (set-members s))) too

```
(define (set-union s1 s2)
  (foldl set-insert s1 (set-members s2)))
```

; And so on. Note that these aren't super efficient

A set module

```
#lang racket
```

```
(provide set set? empty-set? set-members)
```

```
(provide set-contains? set-insert set-union)
```

```
(provide empty-set)
```

```
...
```

Before we continue...

`(struct name field-a field-b...)`

Racket has a very general mechanism for creating structures and the associated procedures

We're not going to use it in this course because it's worth learning how we can do this all ourselves

If you were writing production Racket code, you would definitely want to use this rather than doing it yourself!

Data type with fields

Let's create a type to represent a course

```
(define (course dept num name)
  (list 'course dept num name))
```

```
(define (course? c)
  (and (list? c) (eq? (first c) 'course)))
```

```
(define (course-dept c)
  (if (course? c)
      (second c)
      (error 'course-dept "~v is not a course" c)))
```

```
(define (course-num c)
  (if (course? c)
      (third c)
      (error 'course-num "~v is not a course" c)))

(define (course-name c)
  (if (course? c)
      (fourth c)
      (error 'course-name "~v is not a course" c)))
```

TreeDatatype.rkt

```
#lang racket
; Definition of tree datatype
(provide tree make-tree empty-tree
          tree? empty-tree? leaf?
          tree-value tree-children)
(provide empty-tree T1 T2 T3 T4 T5 T6 T7 T8)
```

Tree constructors and a special value

```
; An empty tree is represented by null  
(define empty-tree null)
```

```
; Create a tree with the given value and children.  
(define (tree value children)  
  (list 'tree value children))
```

```
; Convenience constructor  
; (make-tree v c1 c2 ... cn) is equivalent to  
; (tree v (list c1 c2 ... cn))  
(define (make-tree value . children)  
  (tree value children))
```

Recognizers

; Returns #t if t is an empty tree.

```
(define (empty-tree? t)
  (null? t))
```

; Returns #t if t is a tree (either empty or not).

```
(define (tree? t)
  (cond [(empty-tree? t) #t]
        [(list? t) (eq? (first t) 'tree)]
        [else #f]))
```

; Returns #t if the tree is a leaf.

```
(define (leaf? t)
  (cond [(empty-tree? t) #f]
        [(not (tree? t)) (error 'leaf? "~s is not a tree" t)]
        [else (empty? (tree-children t))]))
```

Accessors

; Returns the tree's value. t must be a nonempty tree.

```
(define (tree-value t)
  (cond [(empty-tree? t)
         (error 'tree-value "argument is an empty tree")]
        [(tree? t) (second t)]
        [else (error 'tree-value "~s is not a tree" t)]))
```

; Returns the tree's children. t must be a nonempty tree.

```
(define (tree-children t)
  (cond [(empty-tree? t)
         (error 'tree-children "argument is an empty tree")]
        [(tree? t) (third t)]
        [else (error 'tree-children "~s is not a tree" t)]))
```


Example trees

```
(define T1 (make-tree 50))  
(define T2 (make-tree 22))  
(define T3 (make-tree 10))  
(define T4 (make-tree 5))  
(define T5 (make-tree 17))  
(define T6 (make-tree 73 T1 T2 T3))  
(define T7 (make-tree 100 T4 T5))  
(define T8 (make-tree 16 T6 T7))
```

A tree is represented as a list ('tree value children).

If you want to count how many children a particular (nonempty) tree `t` has, what's the best way to do it?

A. `(length (tree-children t))`

B. `(length (third t))`

C. `(length (rest t))`

D. `(length (rest (rest t)))`

E. `(length (caddr t))`



chris

@chrsnj



I have a good backtracking joke.
Um, no, actually it is a bad one.

11:31 AM · Jul 28, 2020



See chris's other Tweets

Backtracking

You've seen backtracking before

Anagram lab in CS 150!

- ▶ oberlin student:
let none disturb
run no bed titles
let us not rebind
trust line on bed
but not red lines
bound in letters
let in; runs to bed

Backtracking

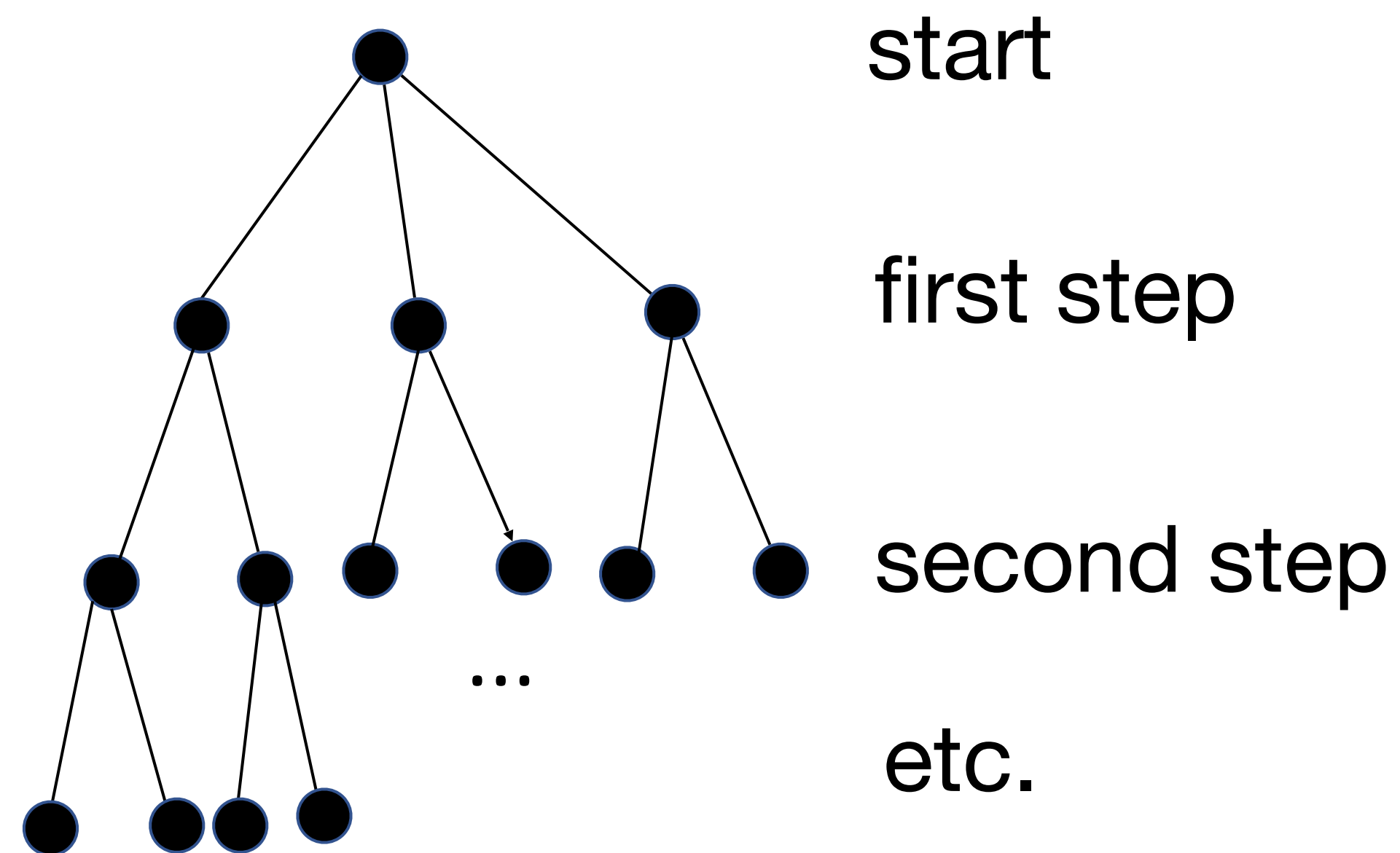
A method to search for all possible elements in the solution space of many problems

- Not efficient: often exponential time
 - Thus it only works on small problems
- + Fairly easy to implement for a wide class of problems

Types of problems

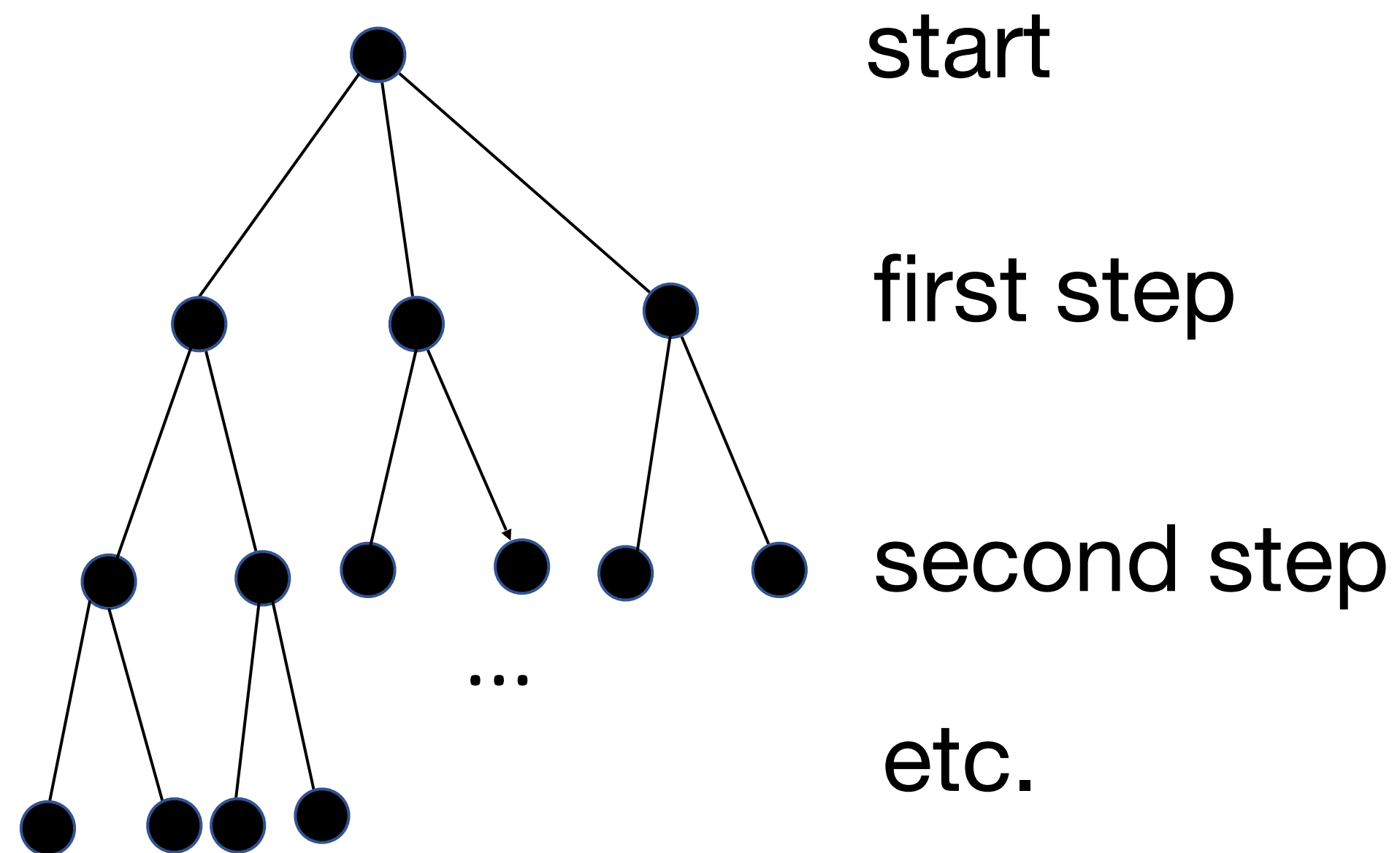
To apply backtracking, the problem needs to have solutions that can be built one step at a time

The solution space for such problems forms a tree



Strategy for solving

- ▶ Choose a step to take
- ▶ If the chosen step cannot possibly lead to a valid solution, back up and make a different choice
- ▶ Repeat this process until a complete solution is found or all possibilities have been exhausted



Examples you've seen before

In the CS 150 Anagrams lab

- ▶ Each step consisted of trying to make a word out of the remaining letters by looking through the words of a dictionary
- ▶ If all letters couldn't be used to make words, you backed up and made different choices

In CS 151, you solved maze using stacks and queues

- ▶ Each step consisted of picking a new cell of the maze to explore
- ▶ If you got stuck, you backed up

n-queens

A famous problem solvable via backtracking

- Place n chess queens on an $n \times n$ chessboard such that no two queens are in the same row, same column, or same diagonal

One step of a solution consists of picking a row for a queen in a given column

So start with the first column, pick a row

Then move to the next column and pick a row

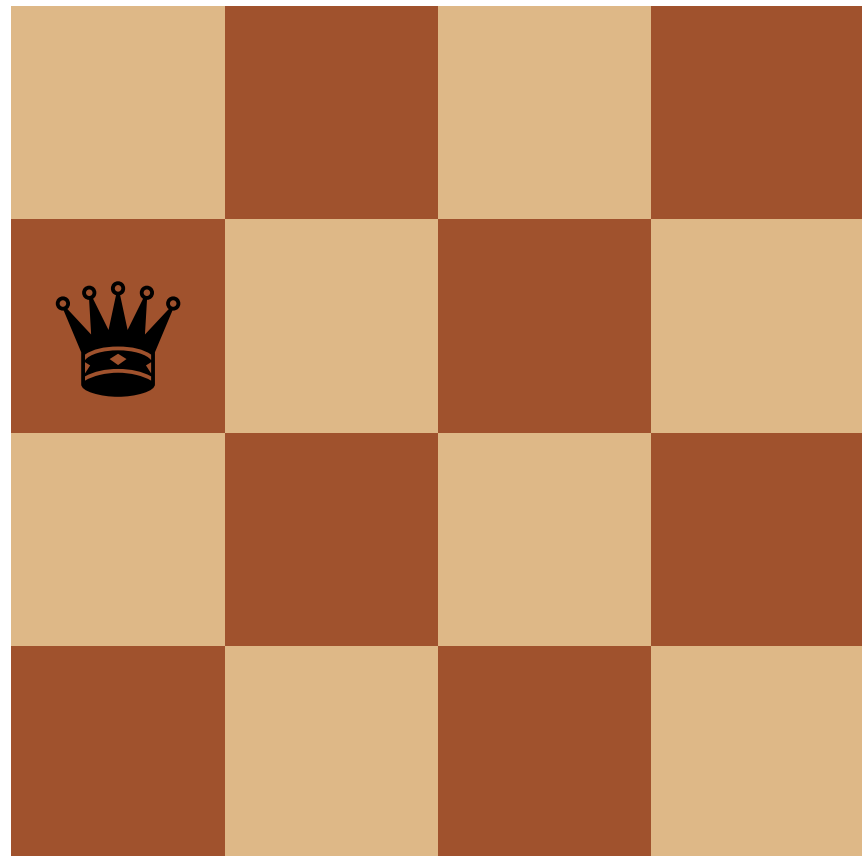
If the partial solution is not valid, backtrack

Repeat until you have a valid solution

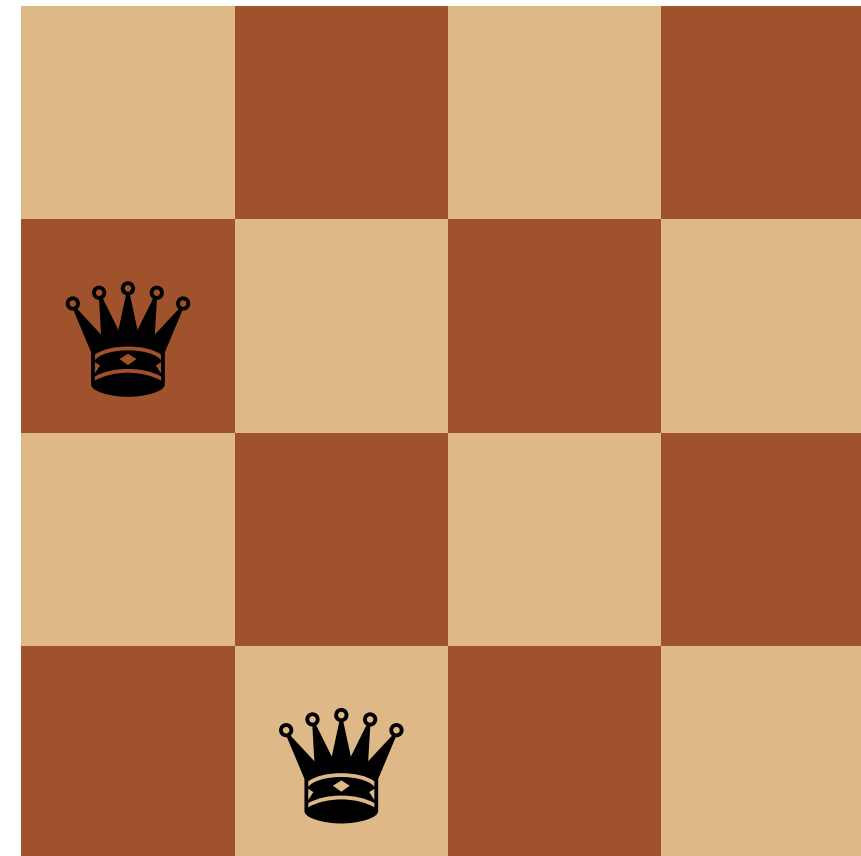
Example: $n = 4$

(Backtracking steps omitted)

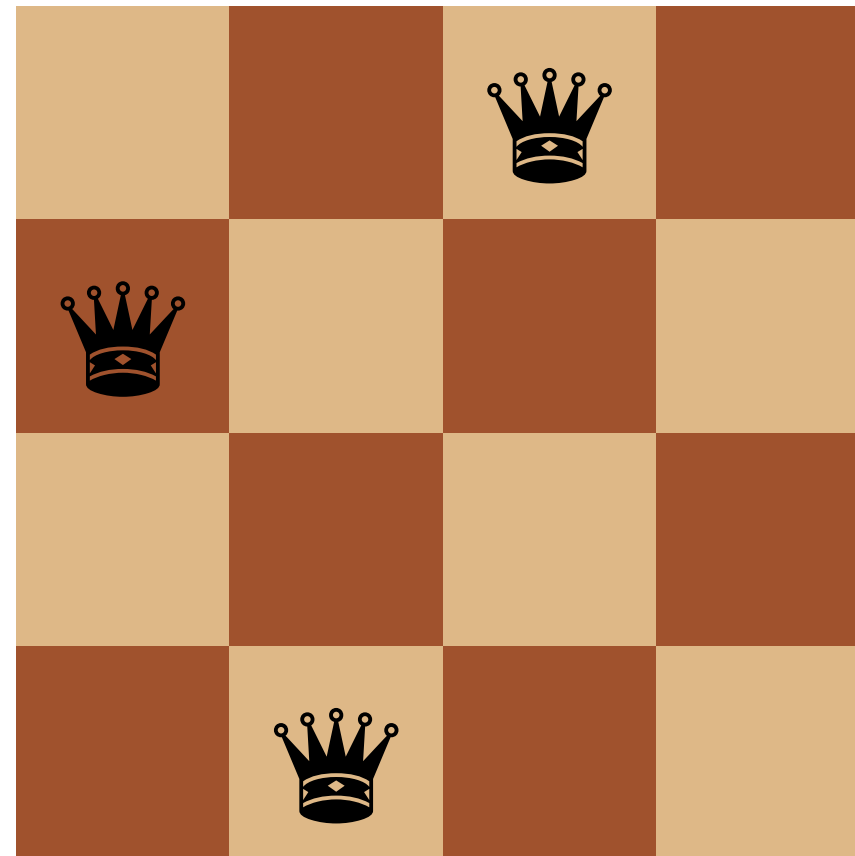
Step 1



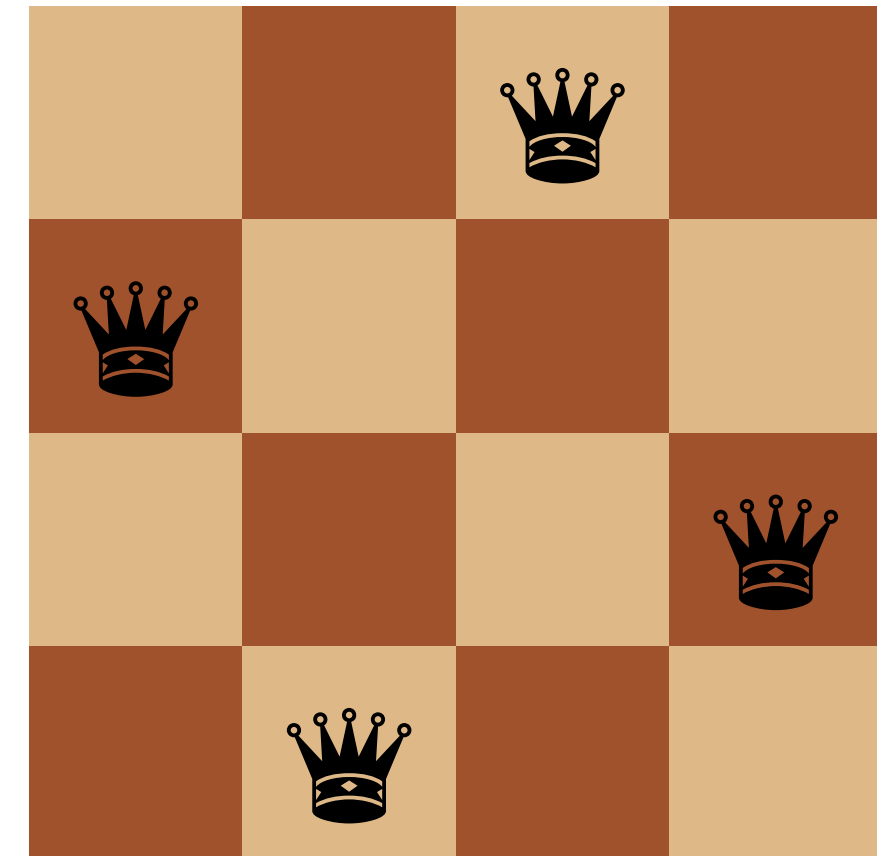
Step 2



Step 3



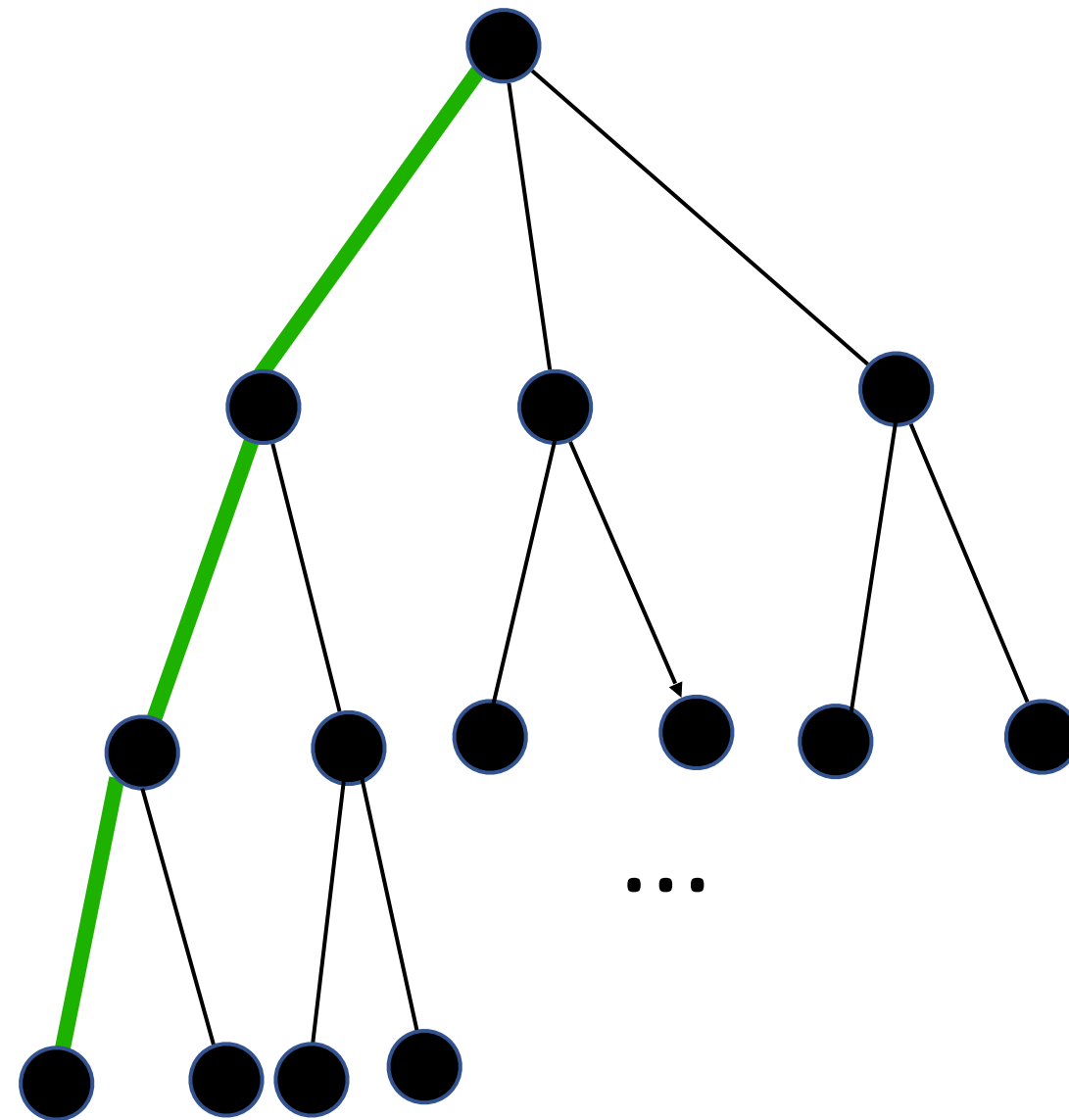
Step 4



Backtracking as search

Backtracking performs a depth-first search through the solution space

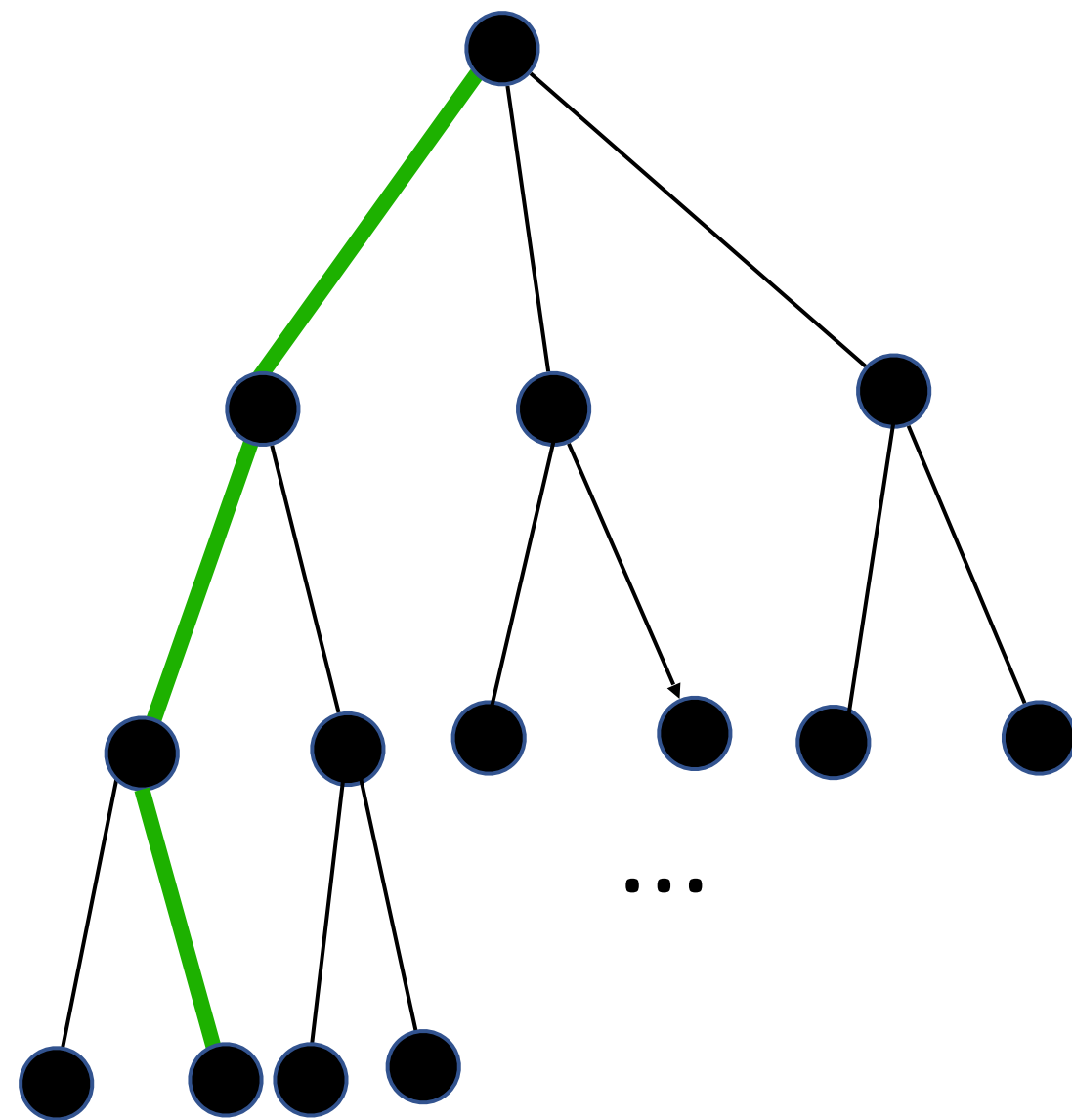
- ▶ It tries the first possible value for the first step
- ▶ Then the first value for the second step
- ▶ And so on



- ▶ If this is a valid solution, we're set!

Backtracking as search

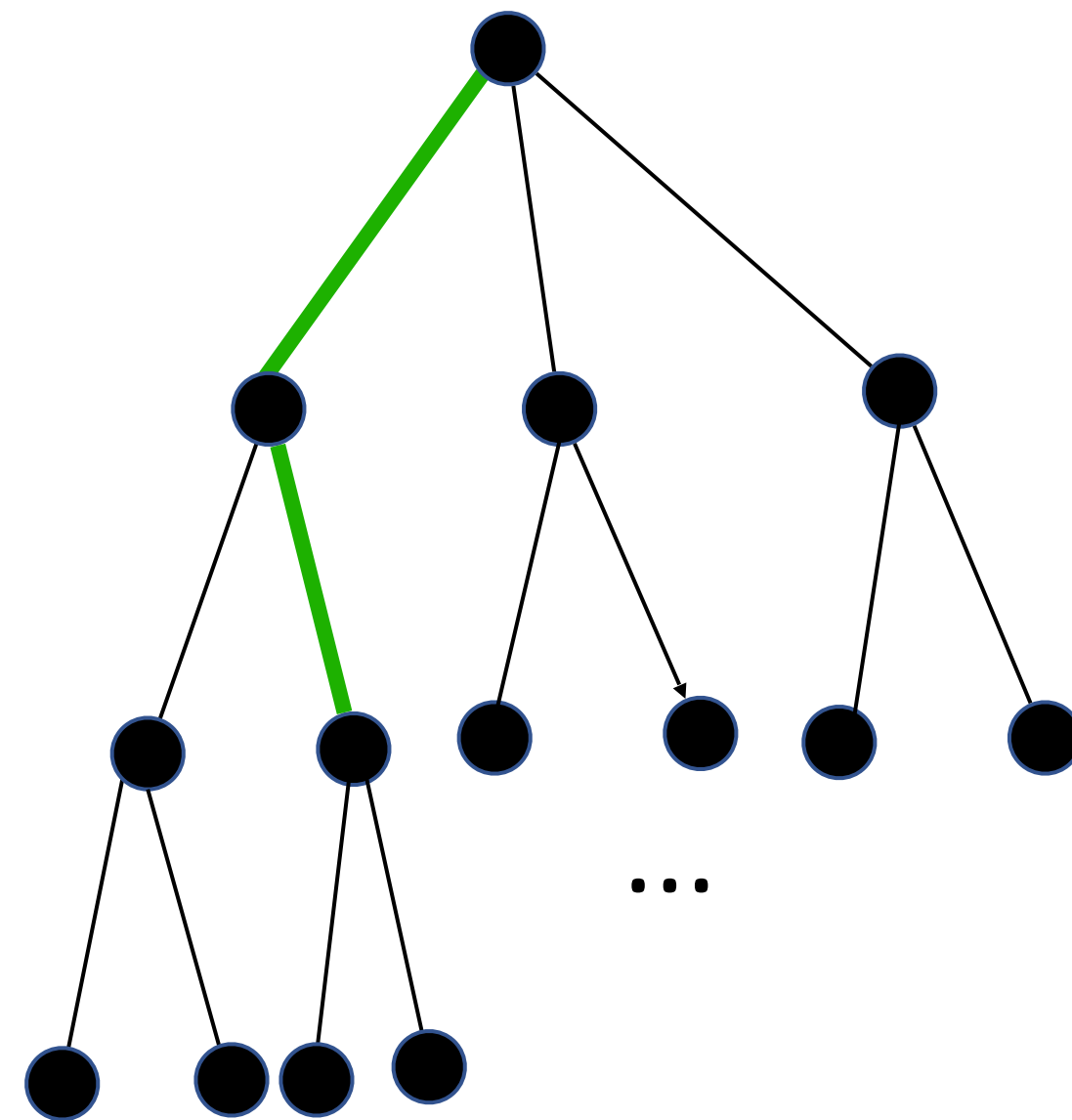
If it's not a valid solution, we back up and make a different choice



Backtracking as search

Suppose this isn't a valid solution so now we're out of options for the third step

We need to make a different second choice



Repeat this until we have a valid solution or none exist

Speeding things up

Backtracking isn't efficient but we can do better than trying every possible value

In many cases, we can test if a partial solution is *feasible*

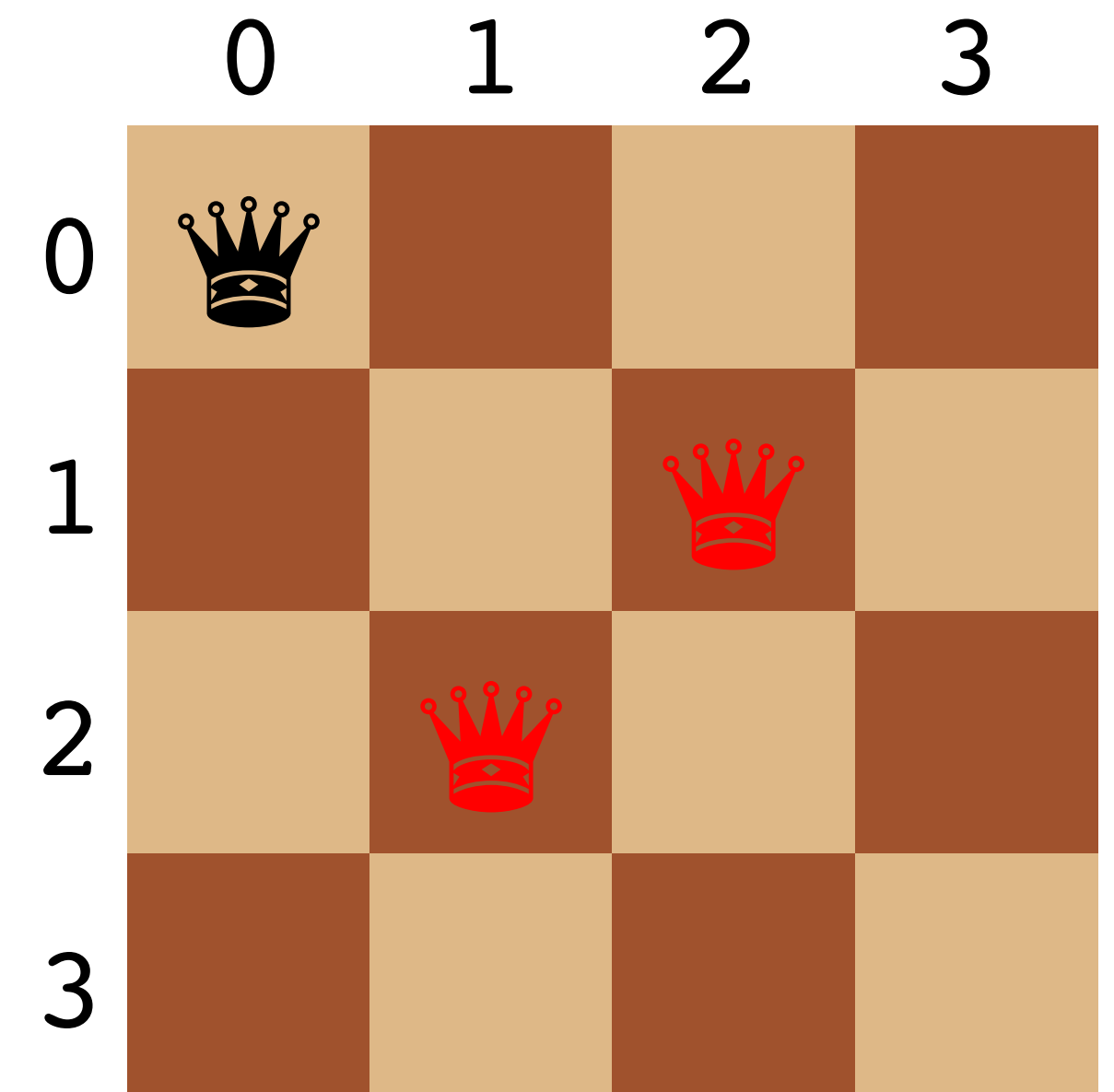
- ▶ If so, continue as before
- ▶ If not, move on to the next (or backtrack) immediately rather than waiting until the whole subtree has been explored

We can do this with n-queens

- ▶ As soon as a partial solution contains two queens in the same row or on the same diagonal, it moves on to the next choice or backtracks

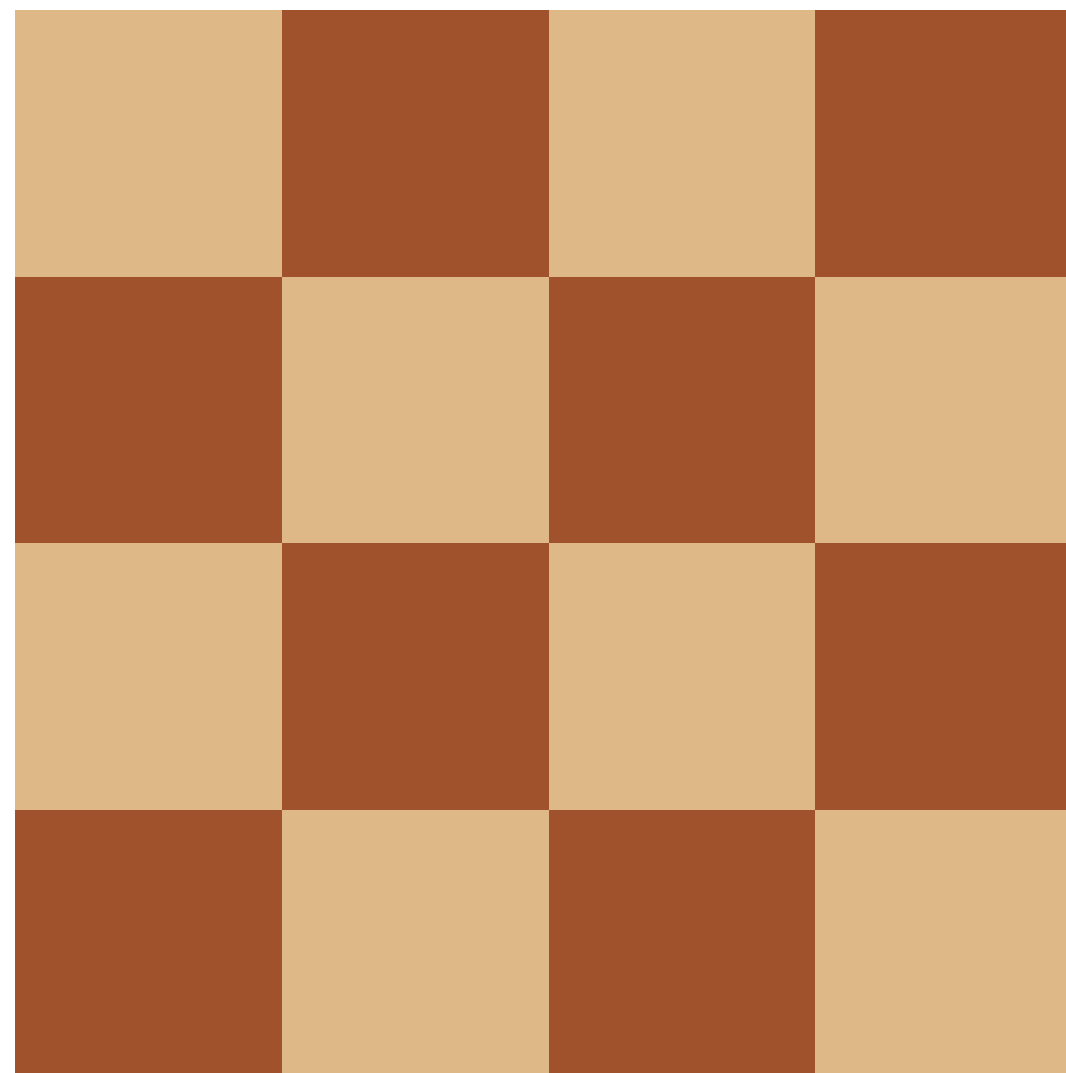
Imagine we're solving 4-queens by picking a row for each column in turn. For column 0 (from left to right), we've selected row 0 (from top to bottom), for column 1, we've selected row 2, and for column 2, we've selected row 1.

At this point, we should either pick the next row for column 2 (i.e., row 2) or backtrack rather than picking a row for column 3. Why?



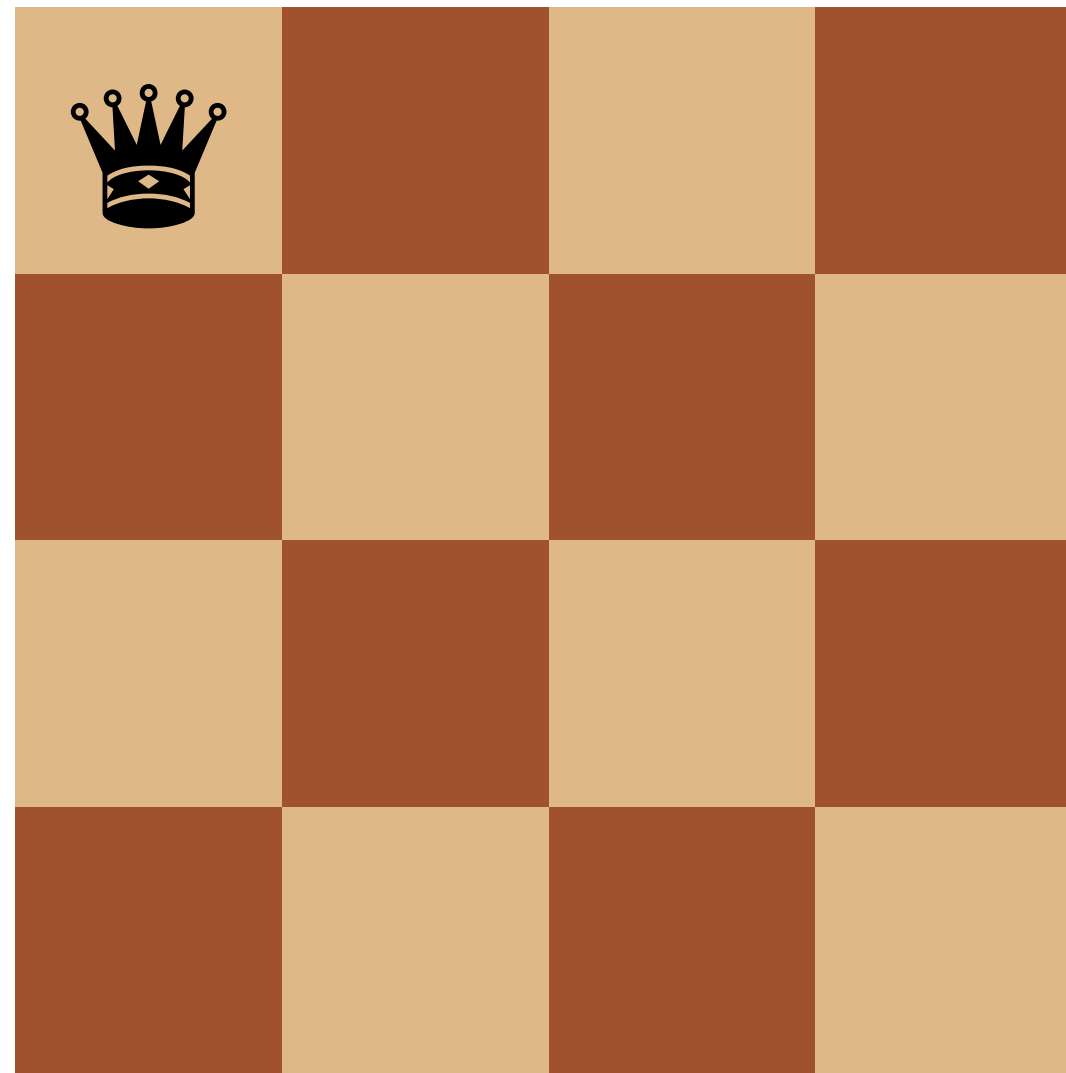
- A. The partial solution is not feasible: no choice for column 3 will be a valid configuration of queens
- B. Queens in columns 1 and 2 share a diagonal
- C. None of other choices for column 2 will work so we need to make a different choice
- D. There's no need to pick the next row or backtrack now; it can do that after picking a row for column 3

Example: $n = 4$



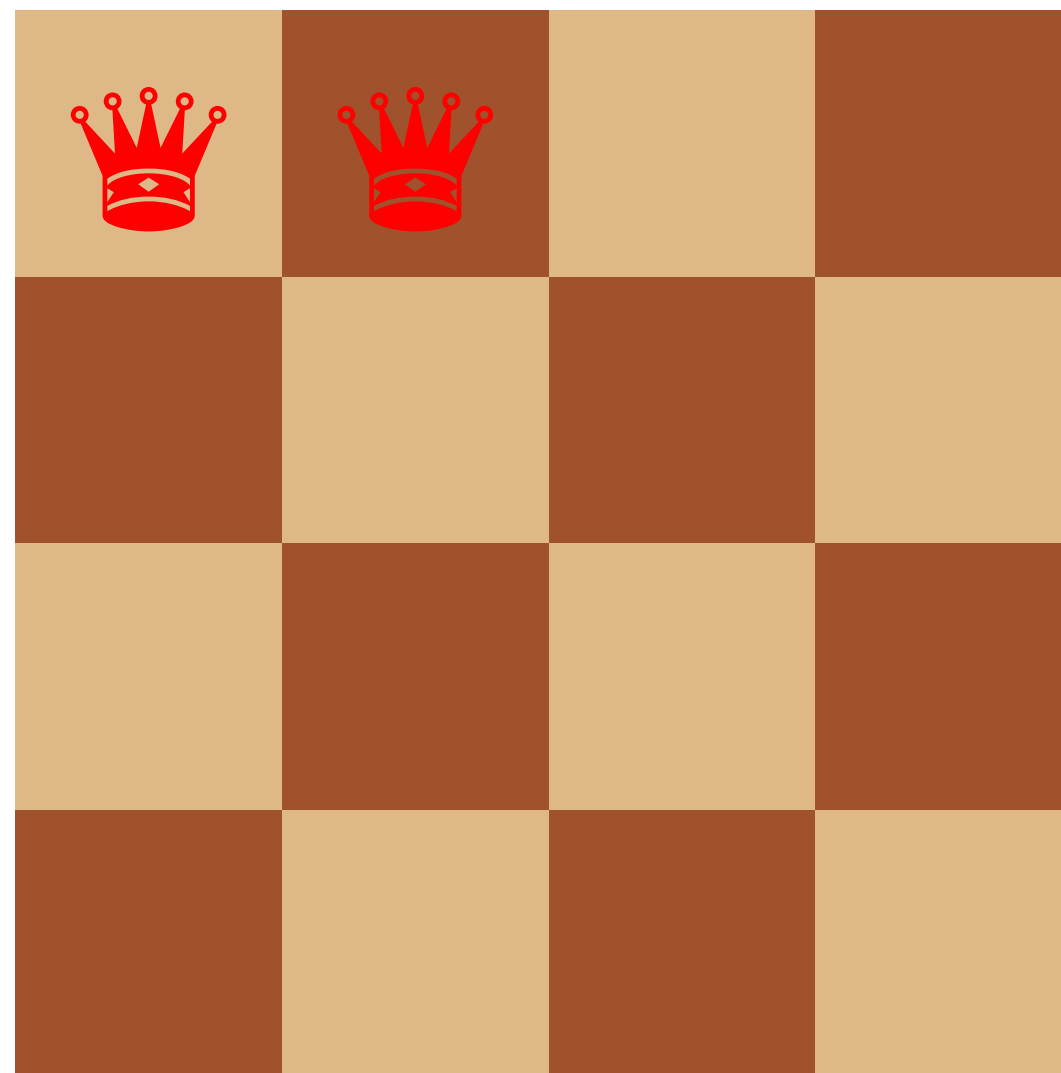
Initial state

Example: $n = 4$



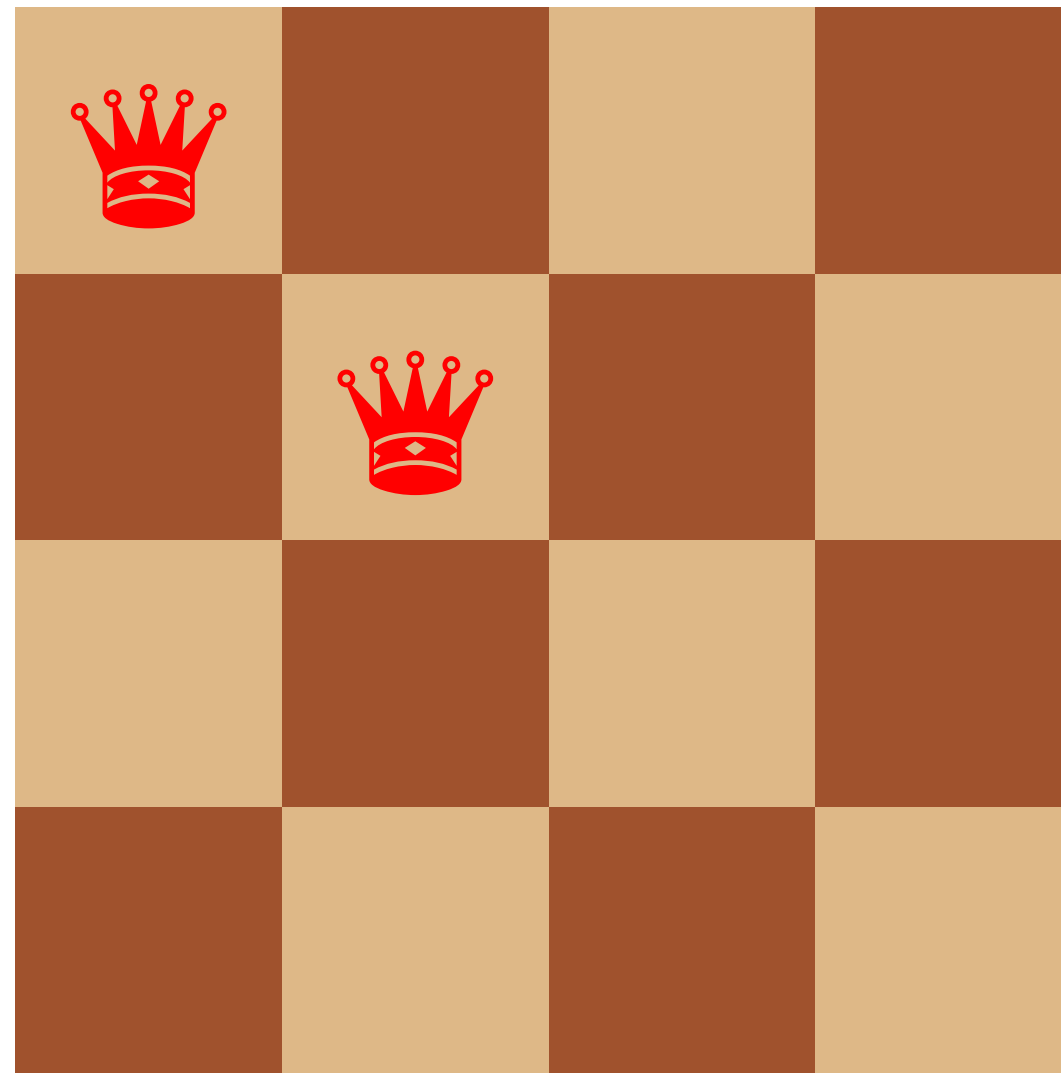
Step 1

Example: $n = 4$



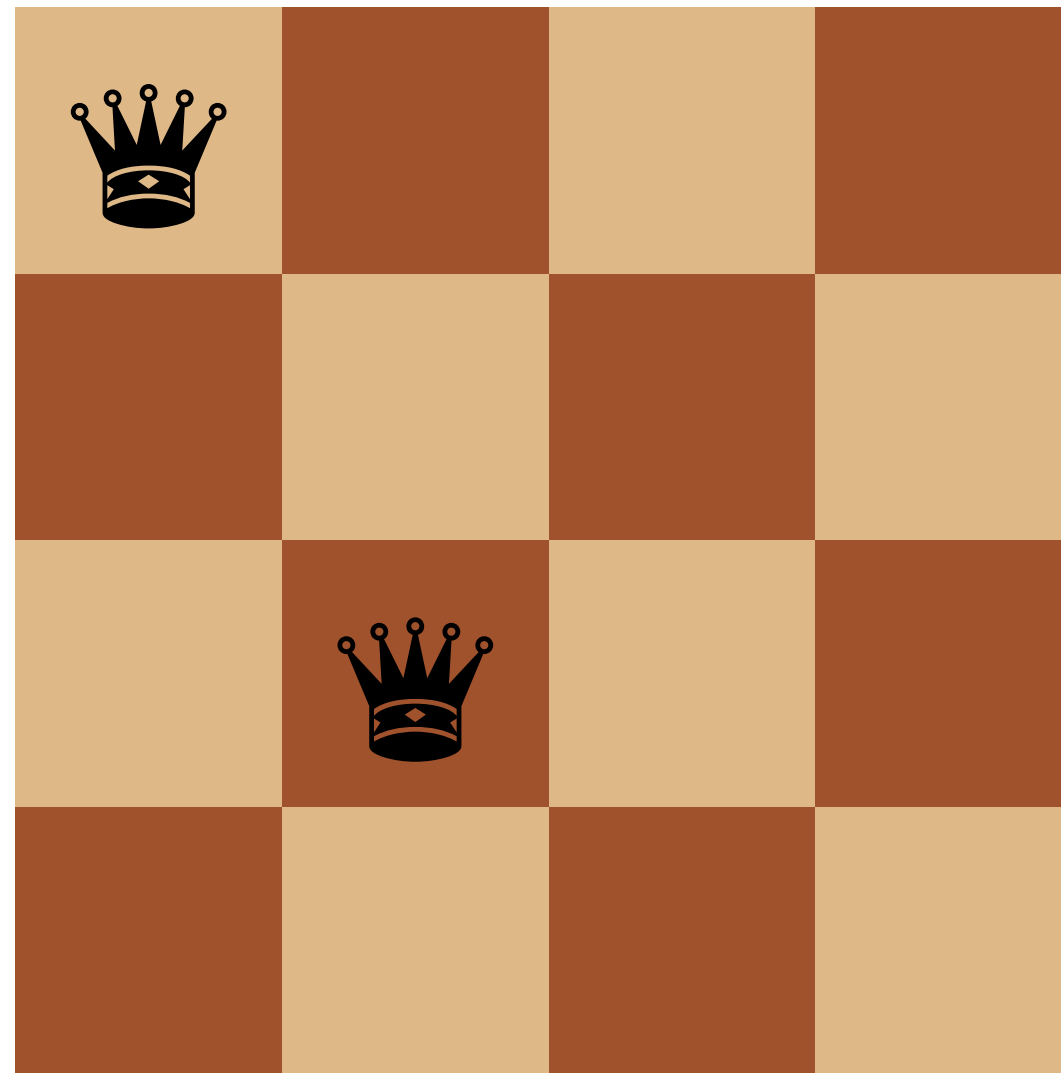
Step 2

Example: $n = 4$



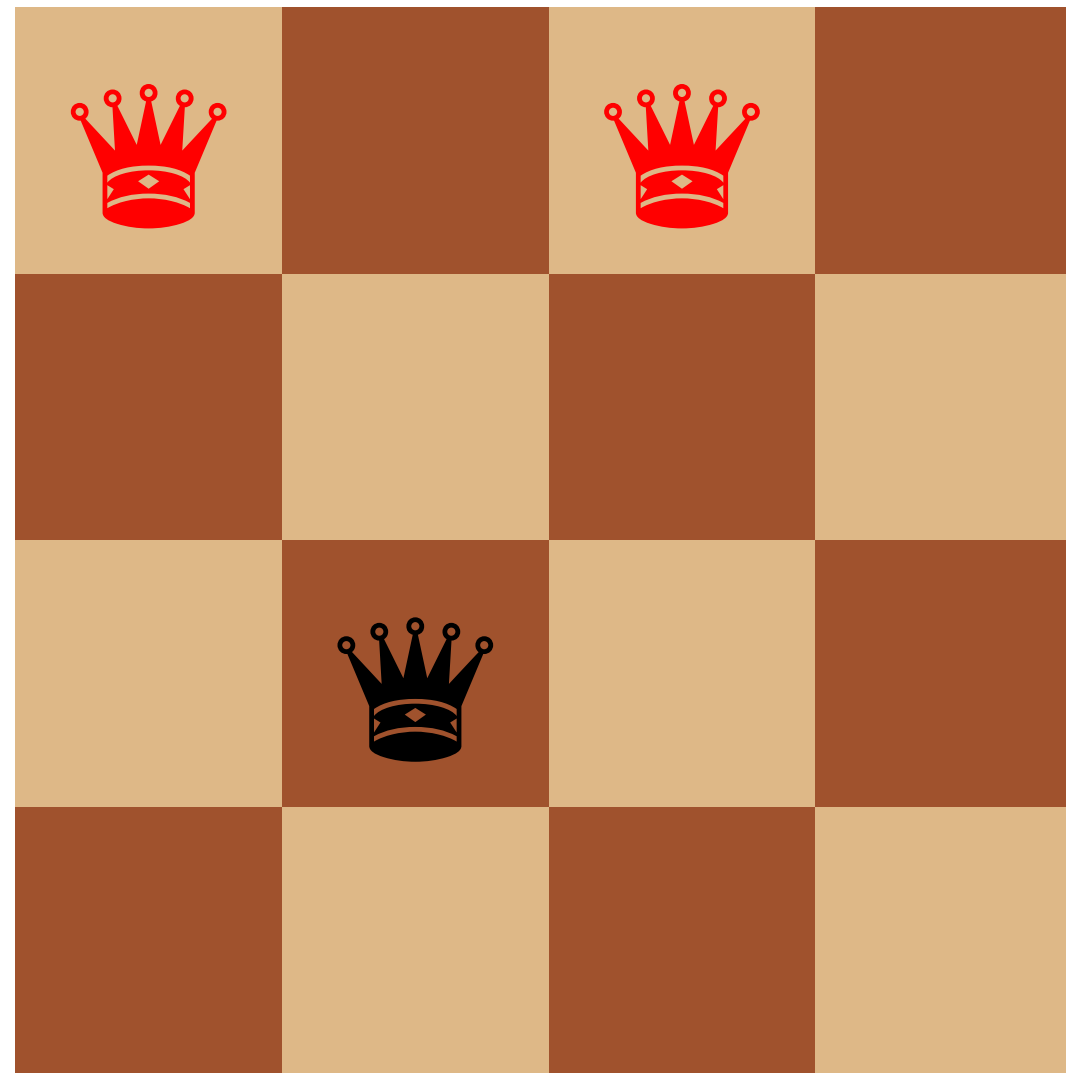
Step 3

Example: $n = 4$



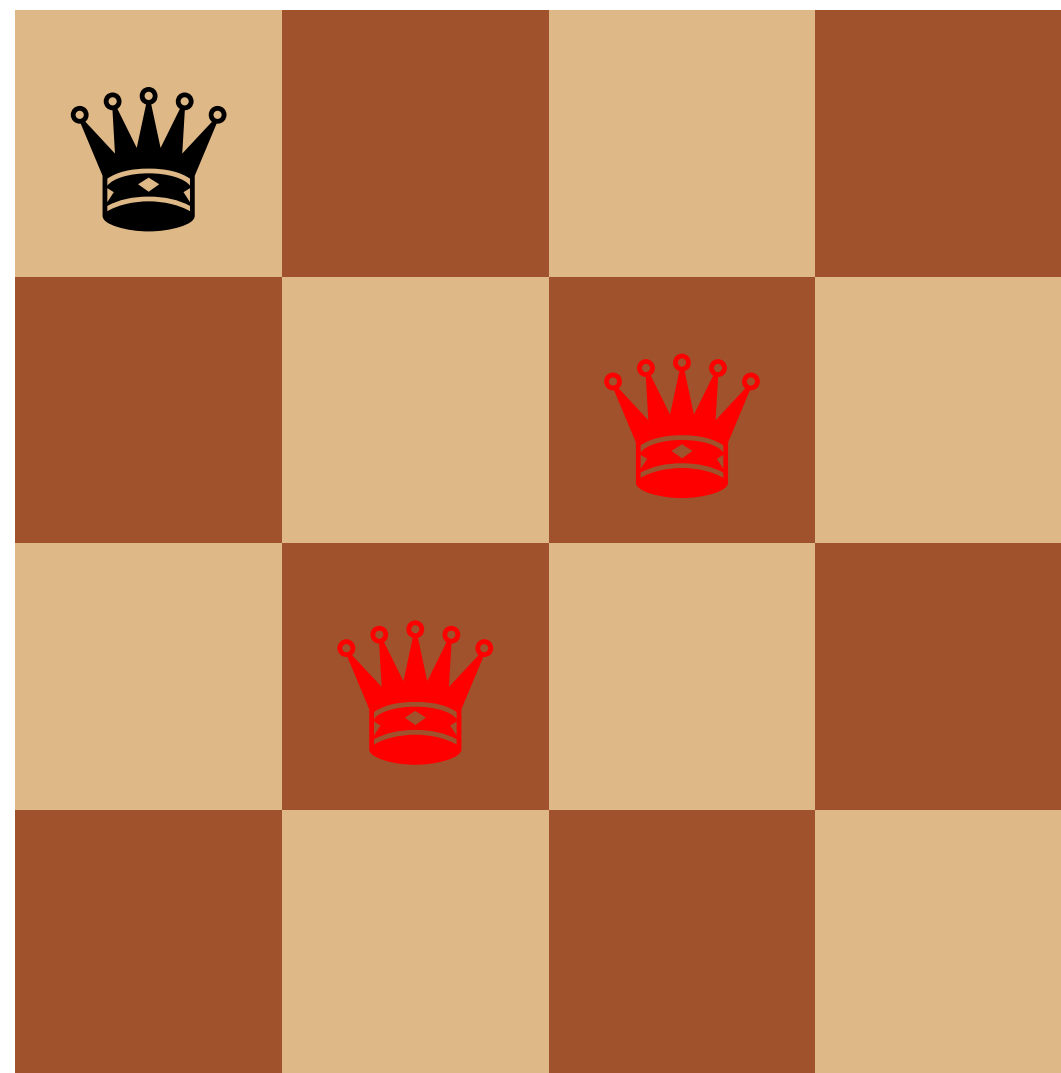
Step 4

Example: $n = 4$



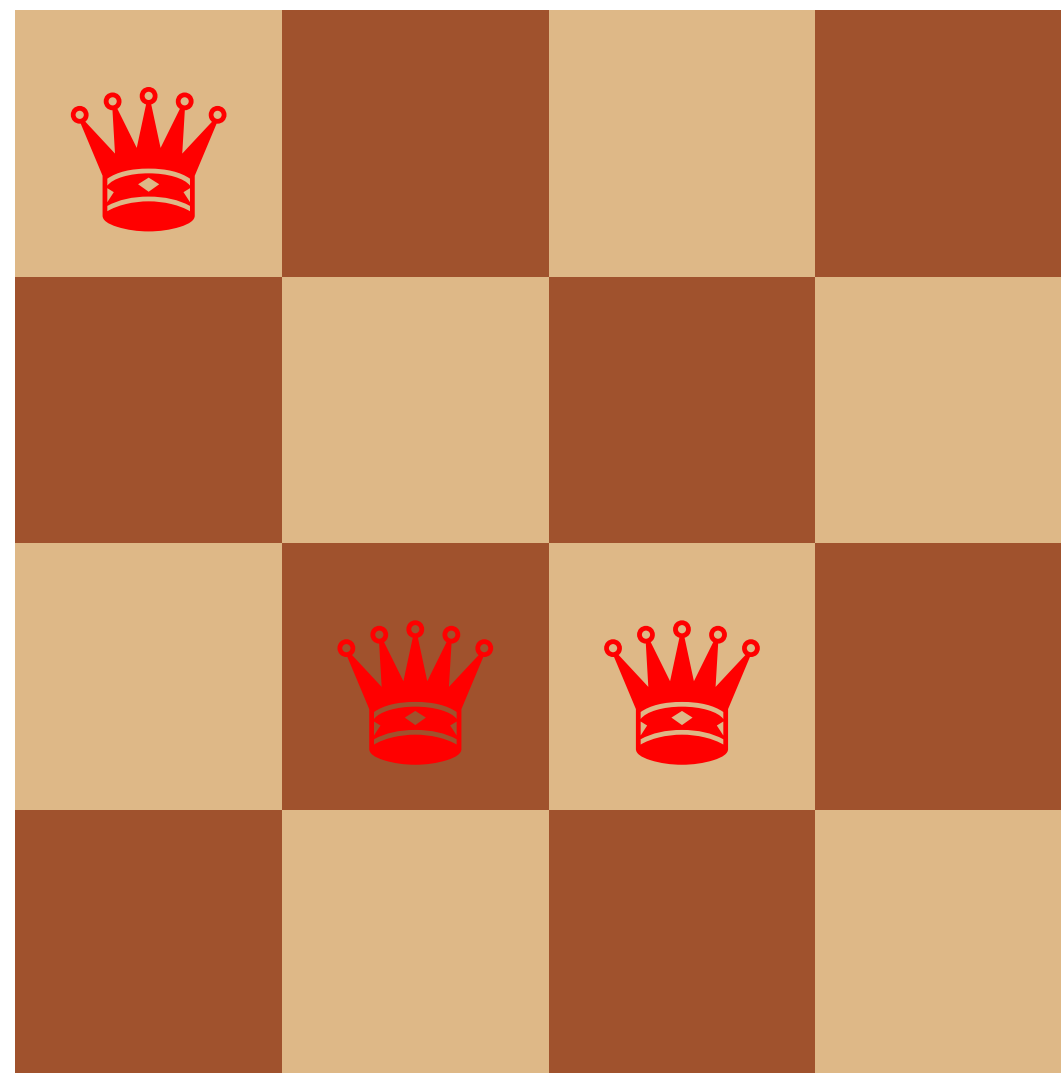
Step 5

Example: $n = 4$



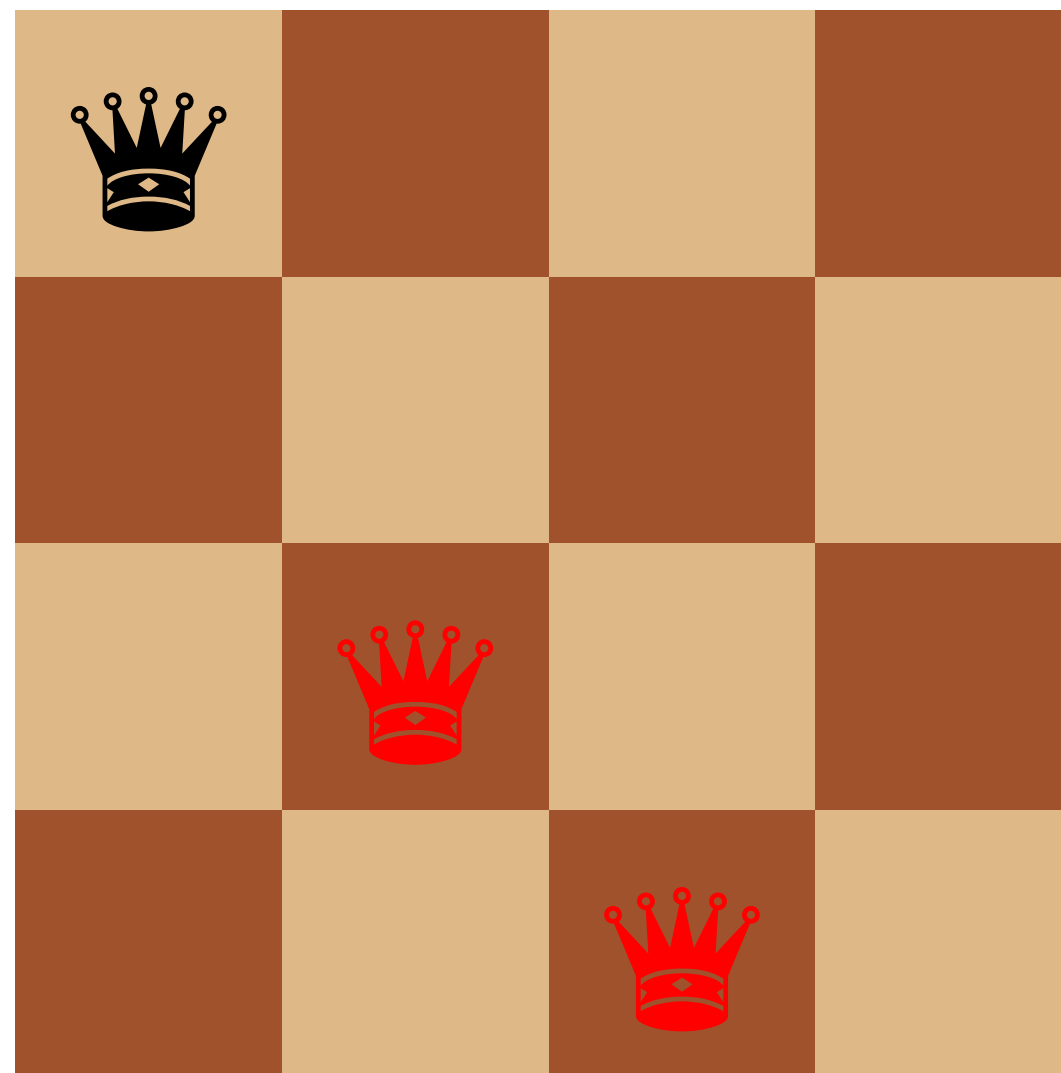
Step 6

Example: $n = 4$



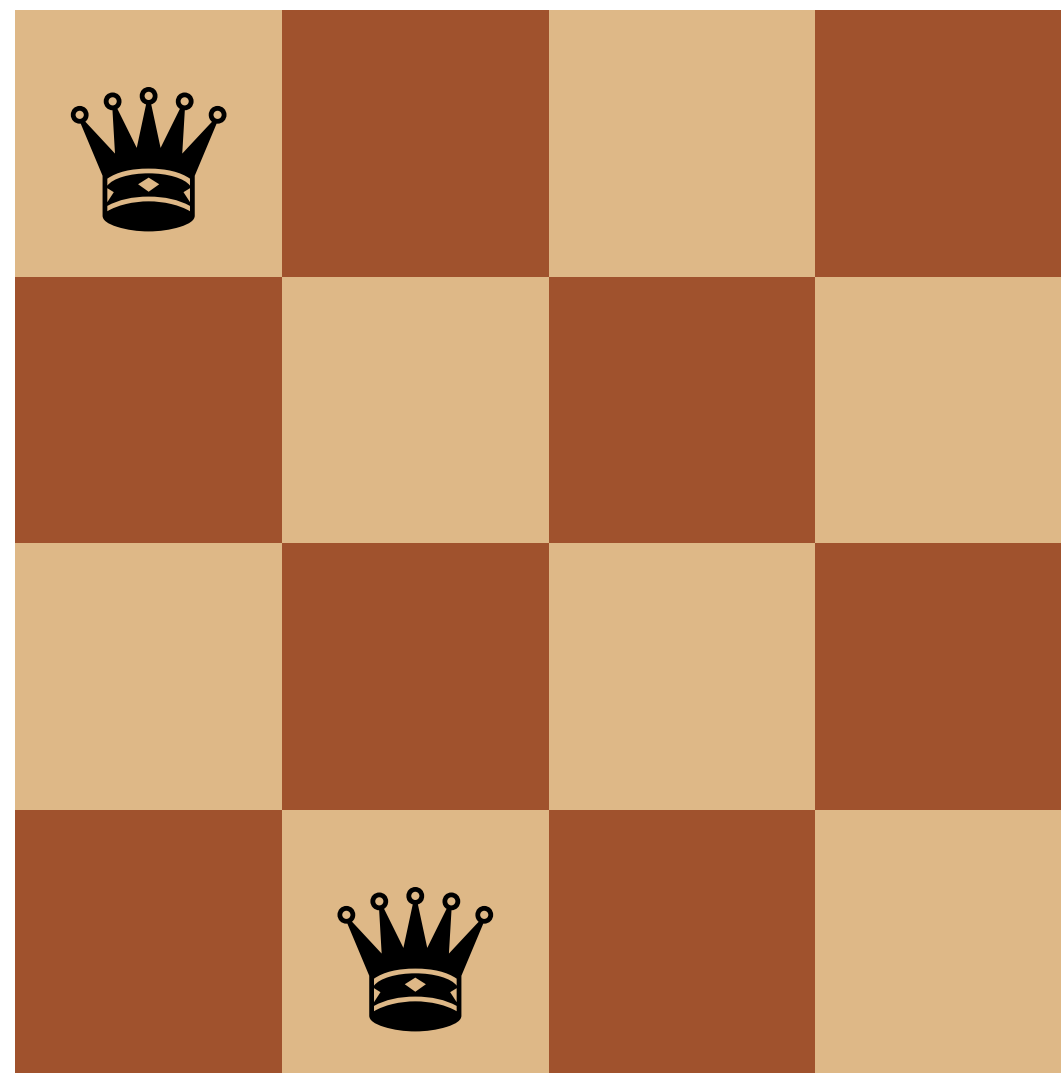
Step 7

Example: $n = 4$



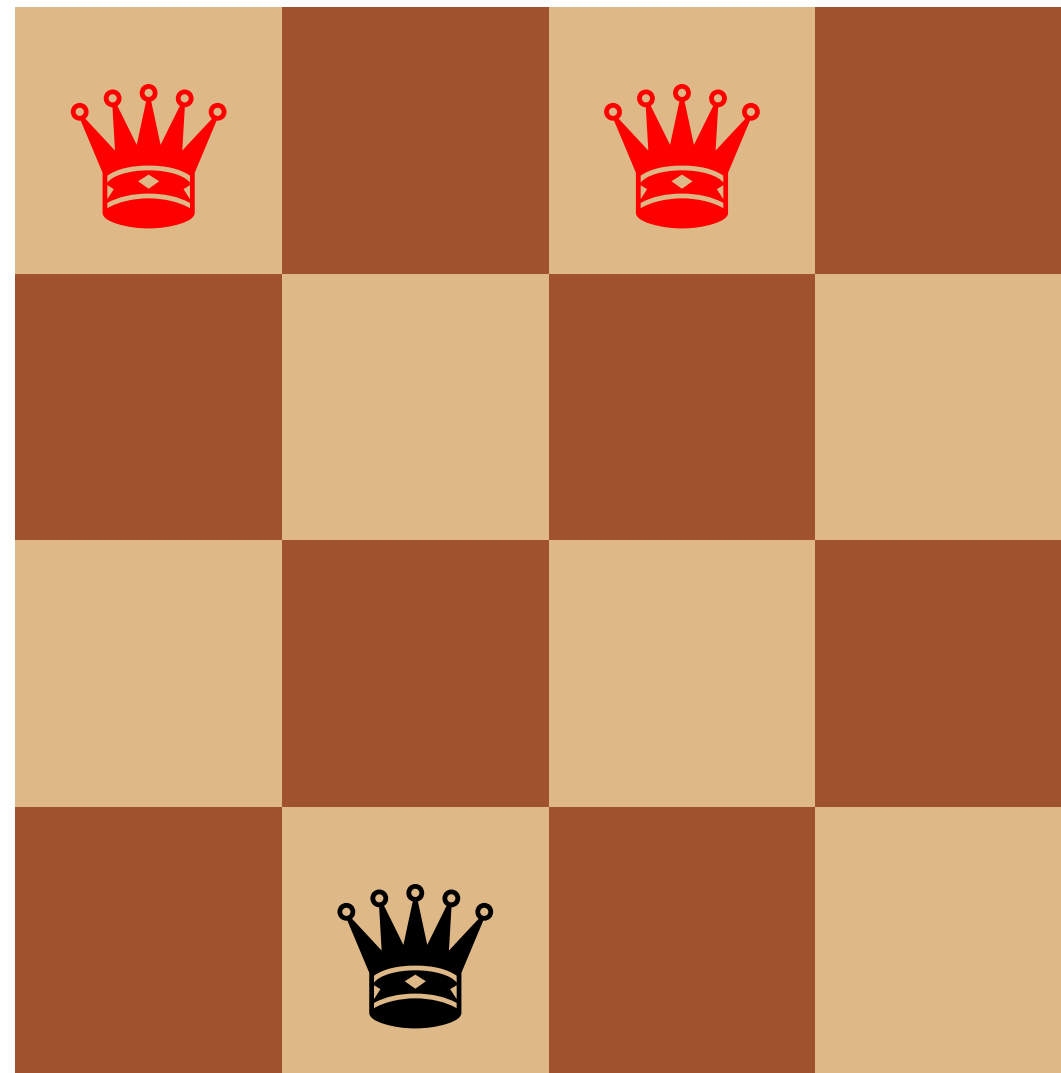
Step 8

Example: $n = 4$



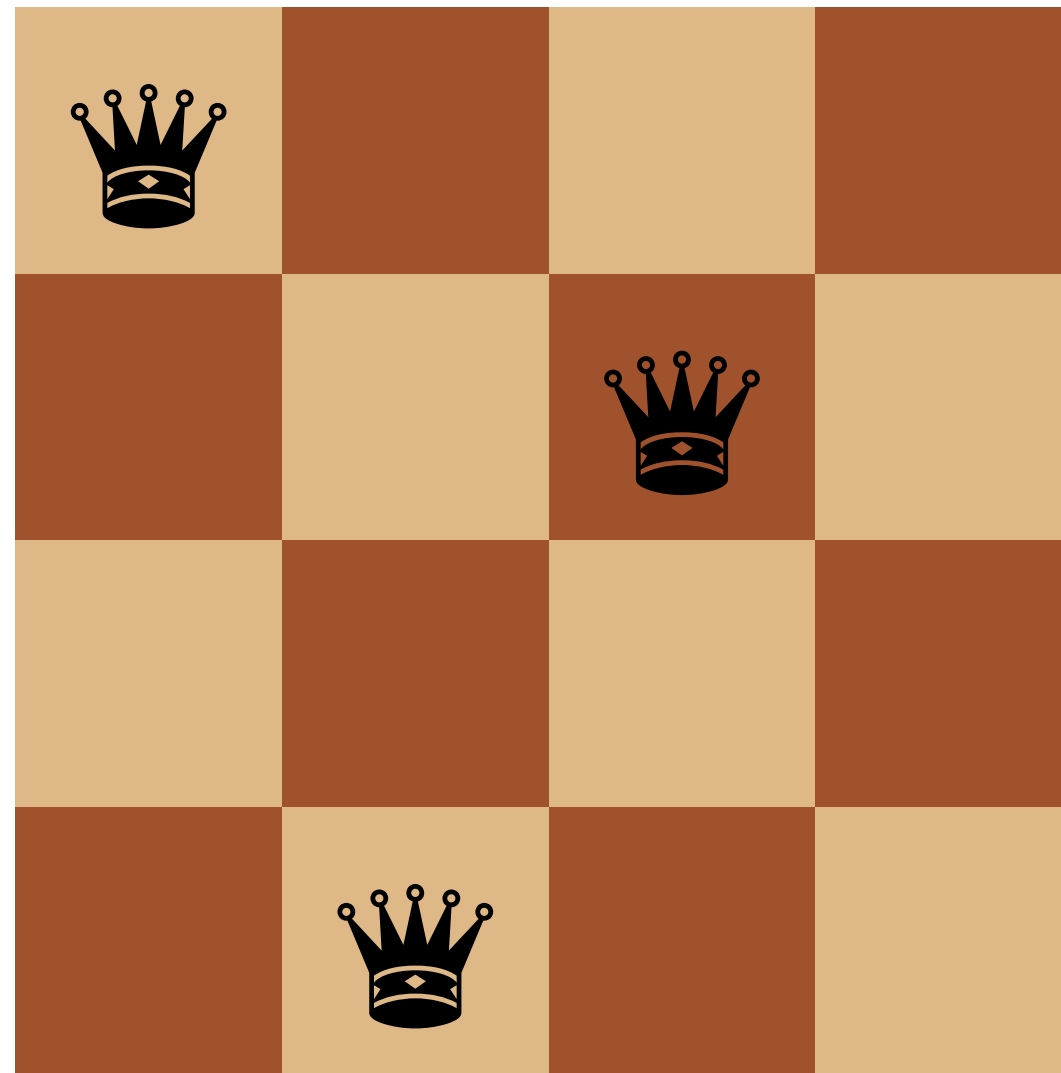
Step 9

Example: $n = 4$



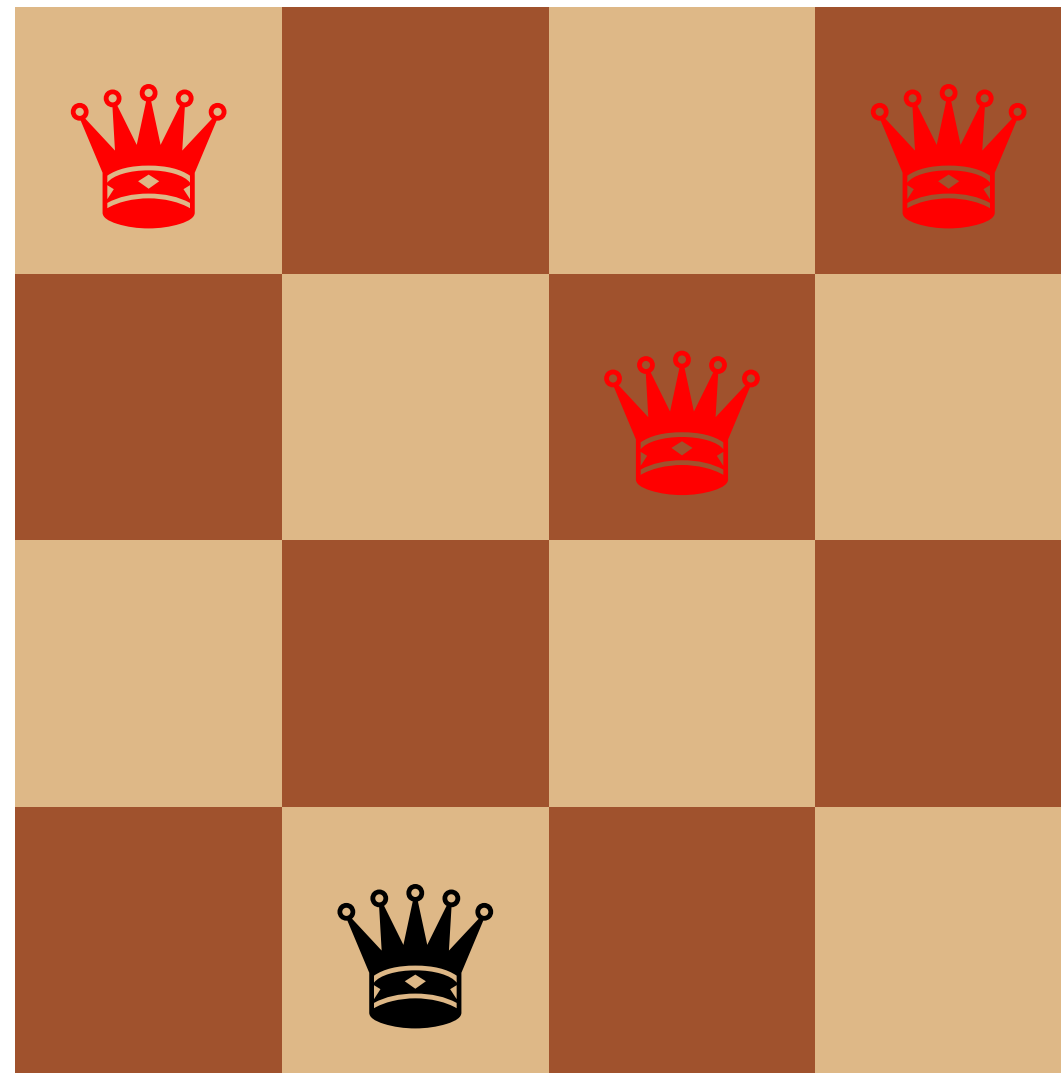
Step 10

Example: $n = 4$



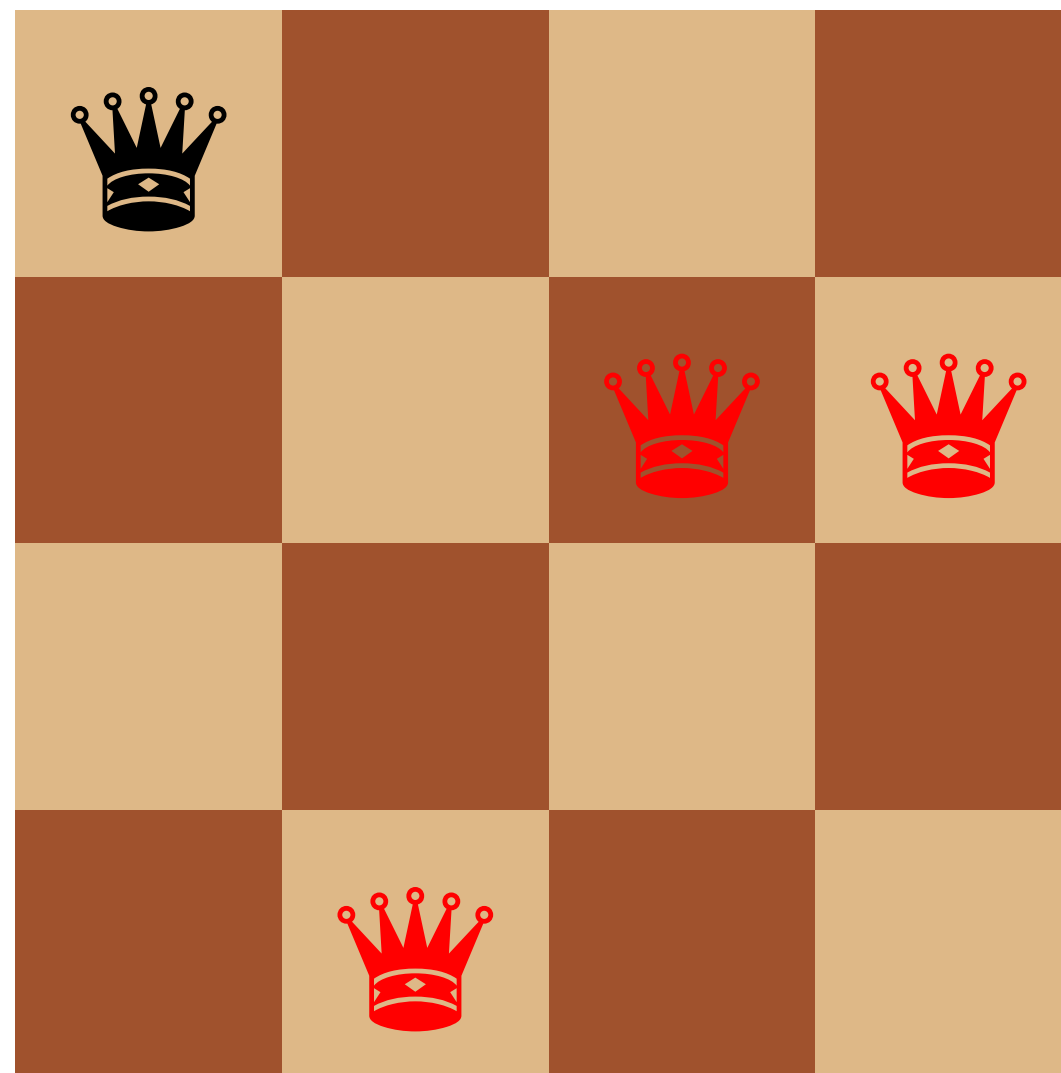
Step 11

Example: $n = 4$



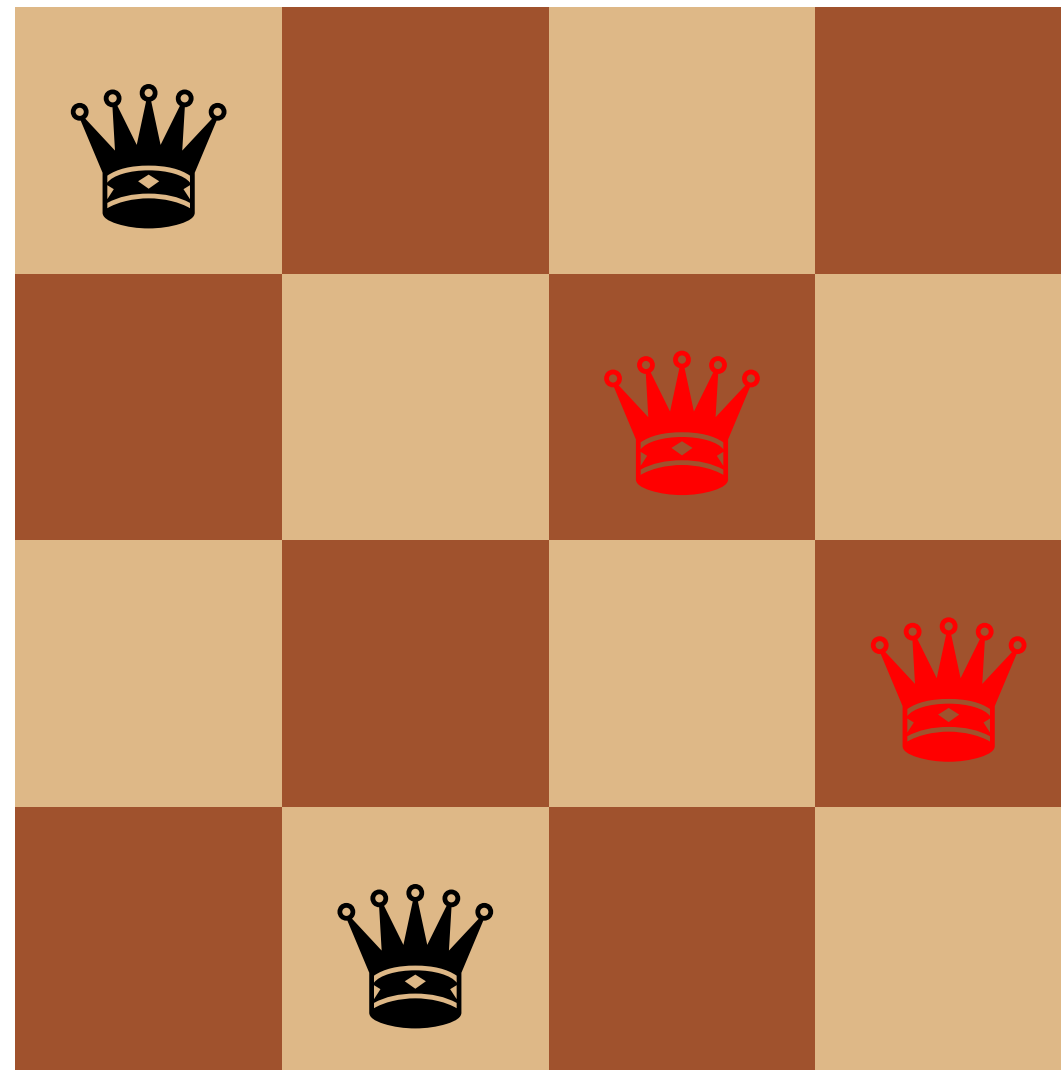
Step 12

Example: $n = 4$



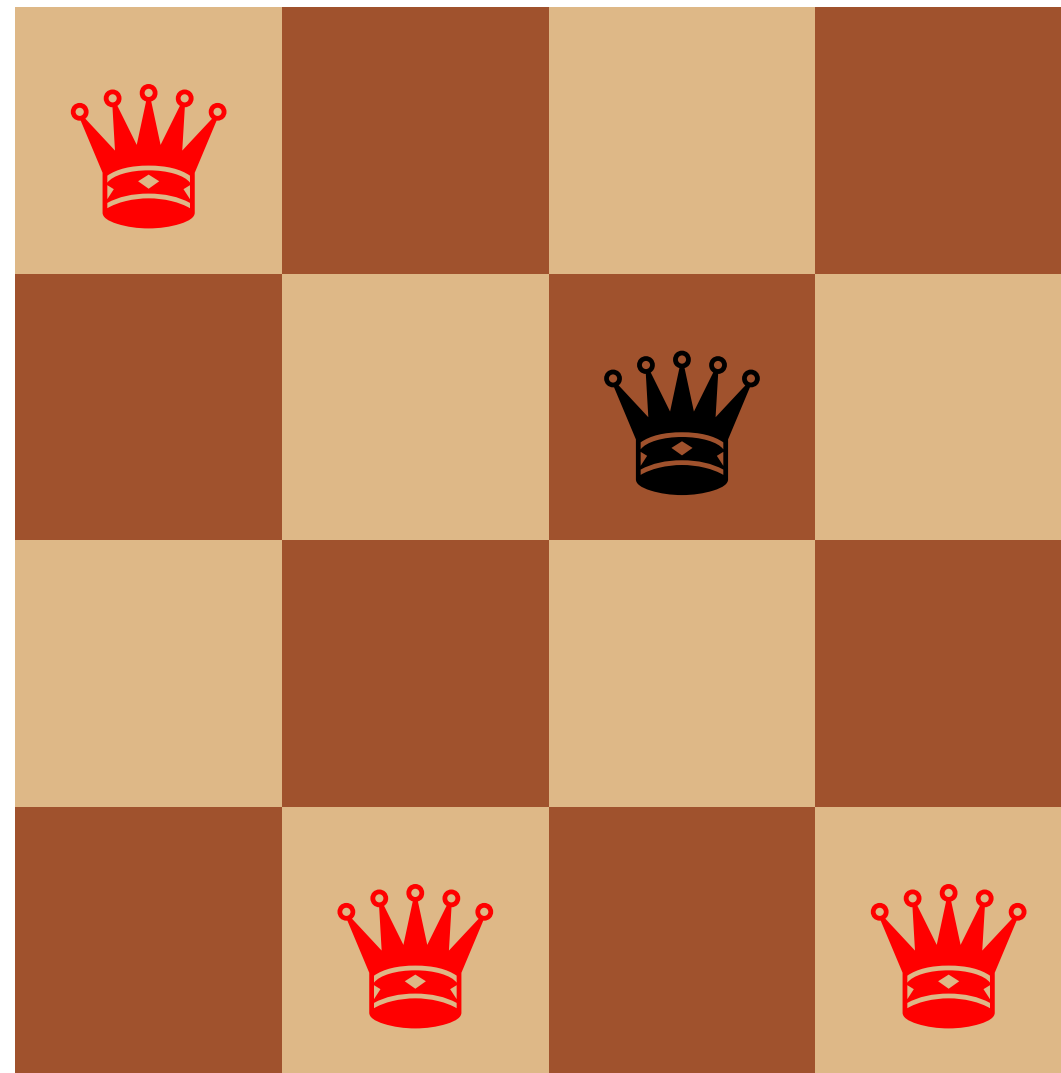
Step 13

Example: $n = 4$



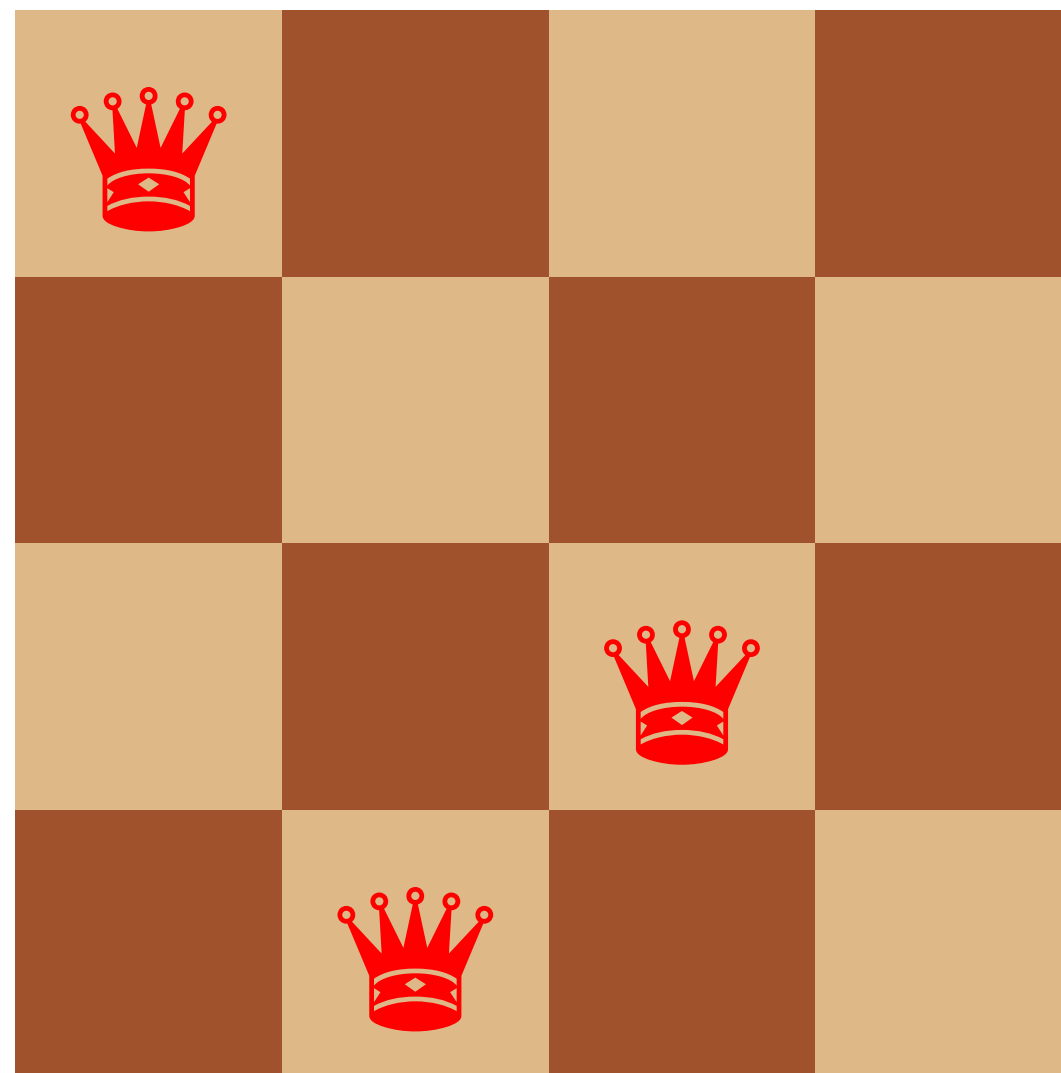
Step 14

Example: $n = 4$



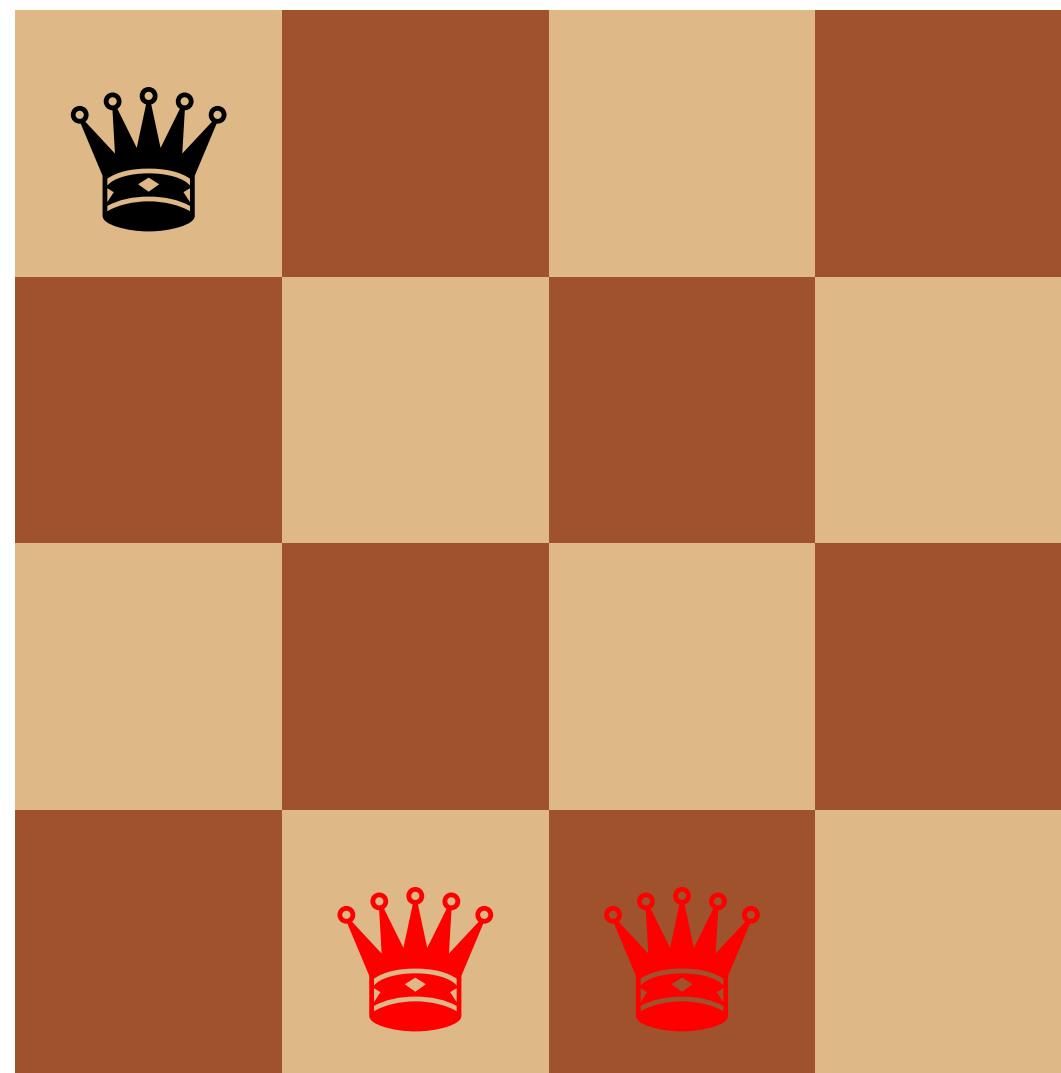
Step 15

Example: $n = 4$



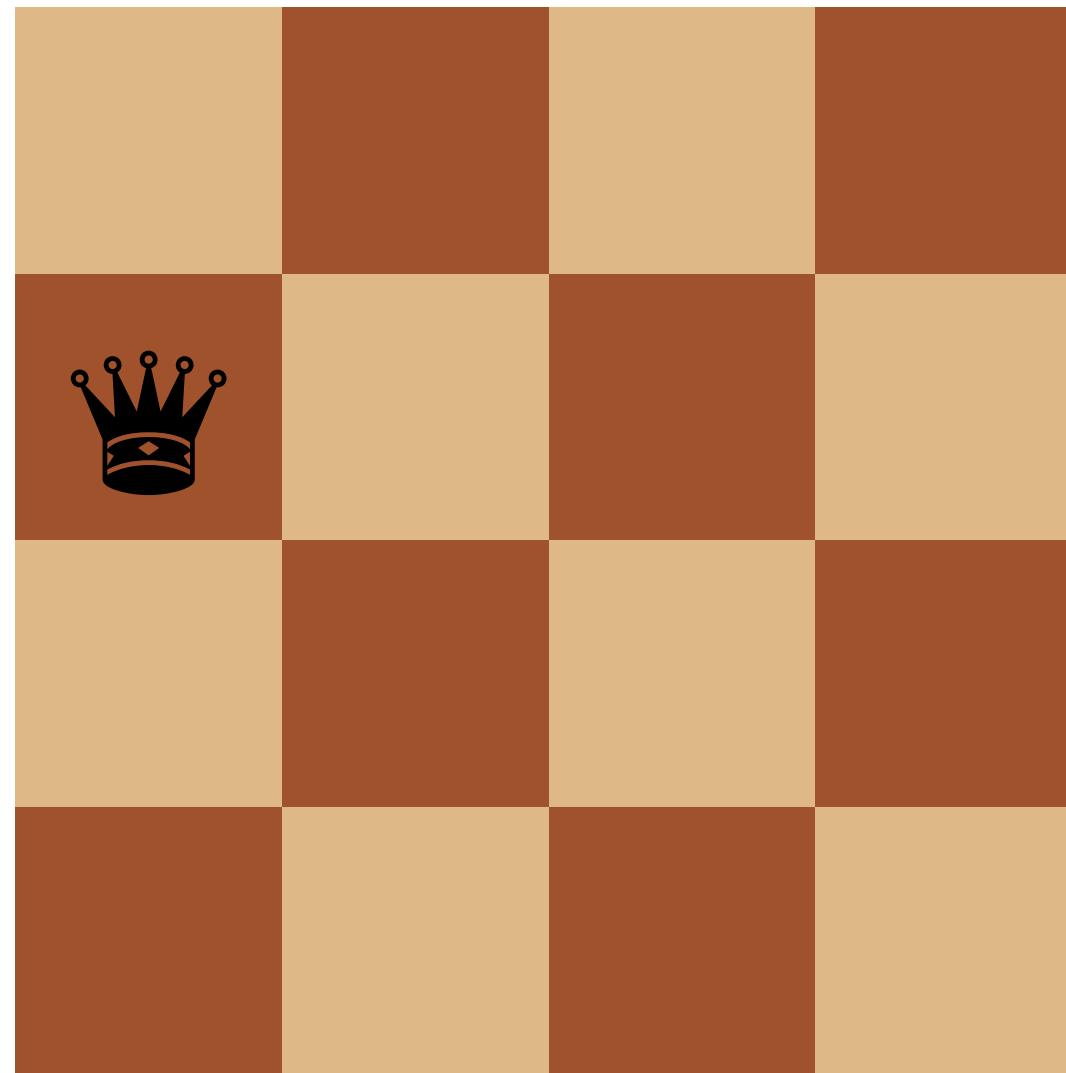
Step 16

Example: $n = 4$



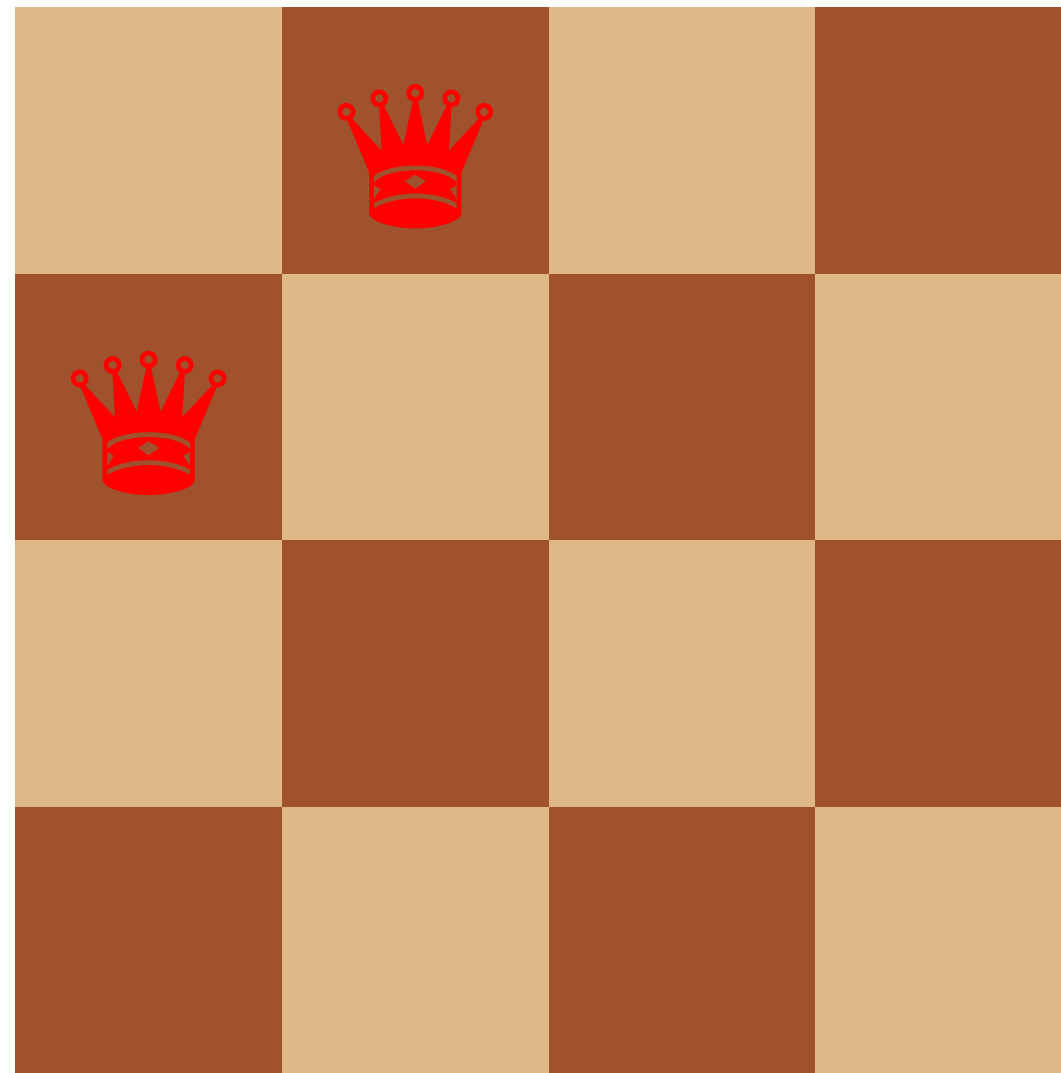
Step 17

Example: $n = 4$



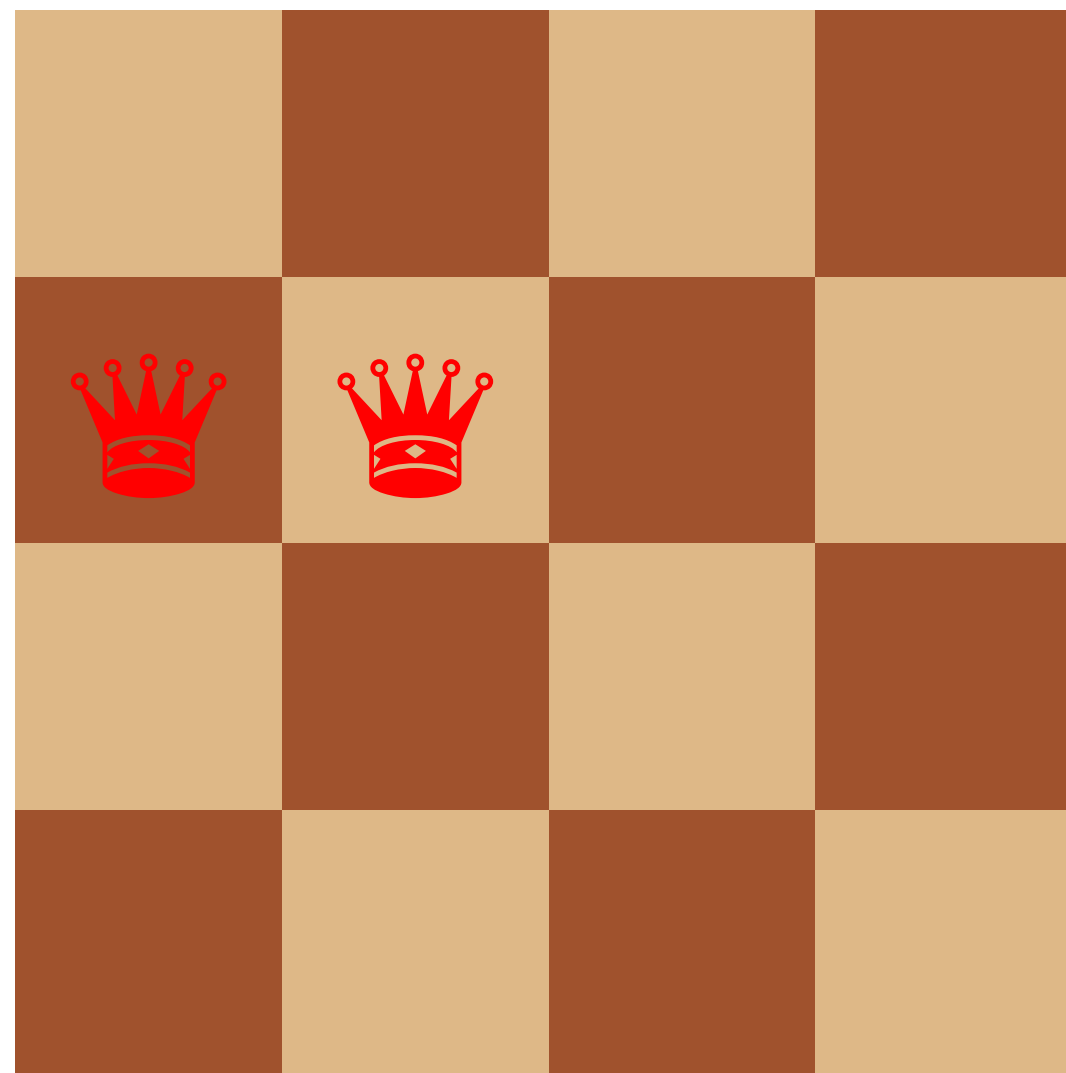
Step 18

Example: $n = 4$



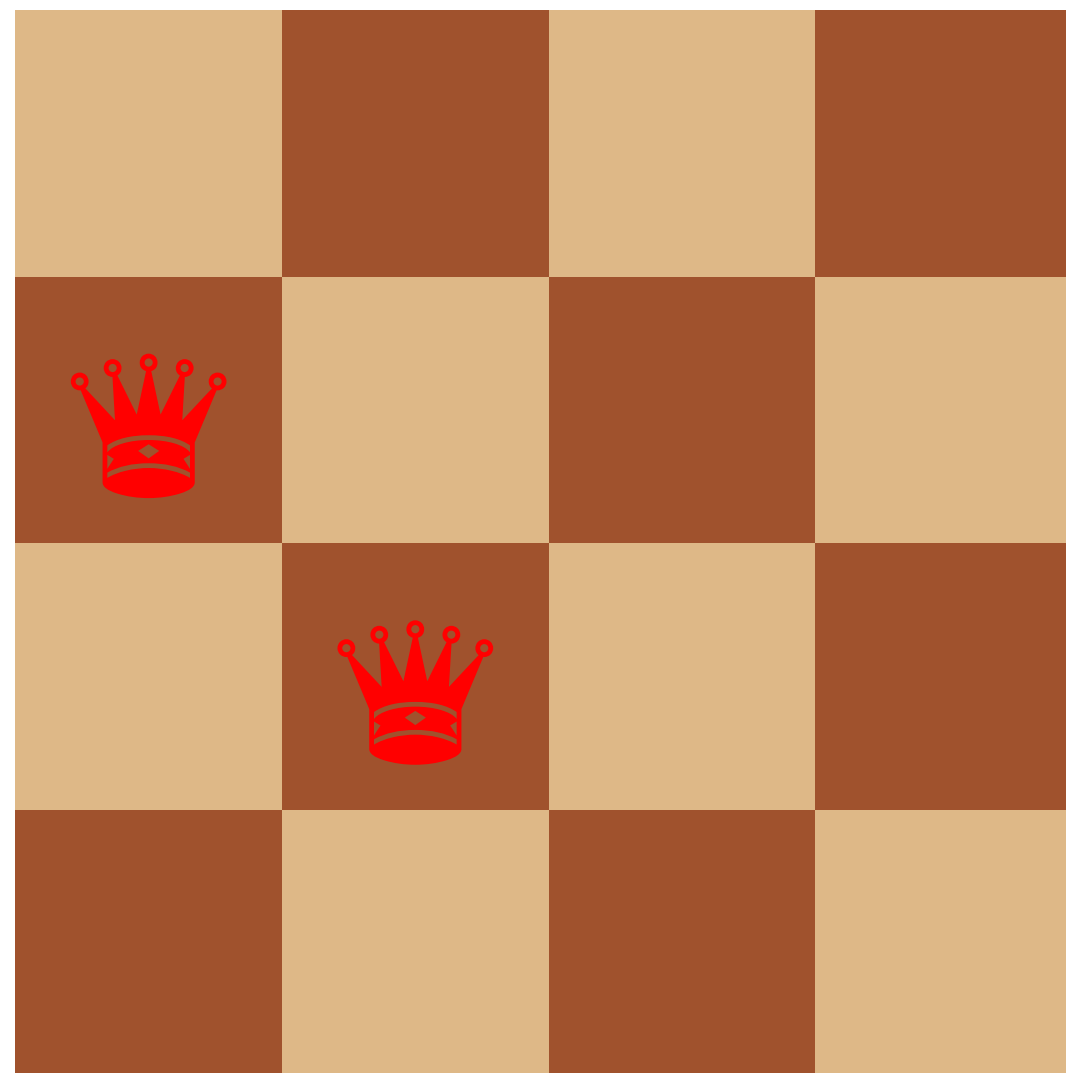
Step 19

Example: $n = 4$



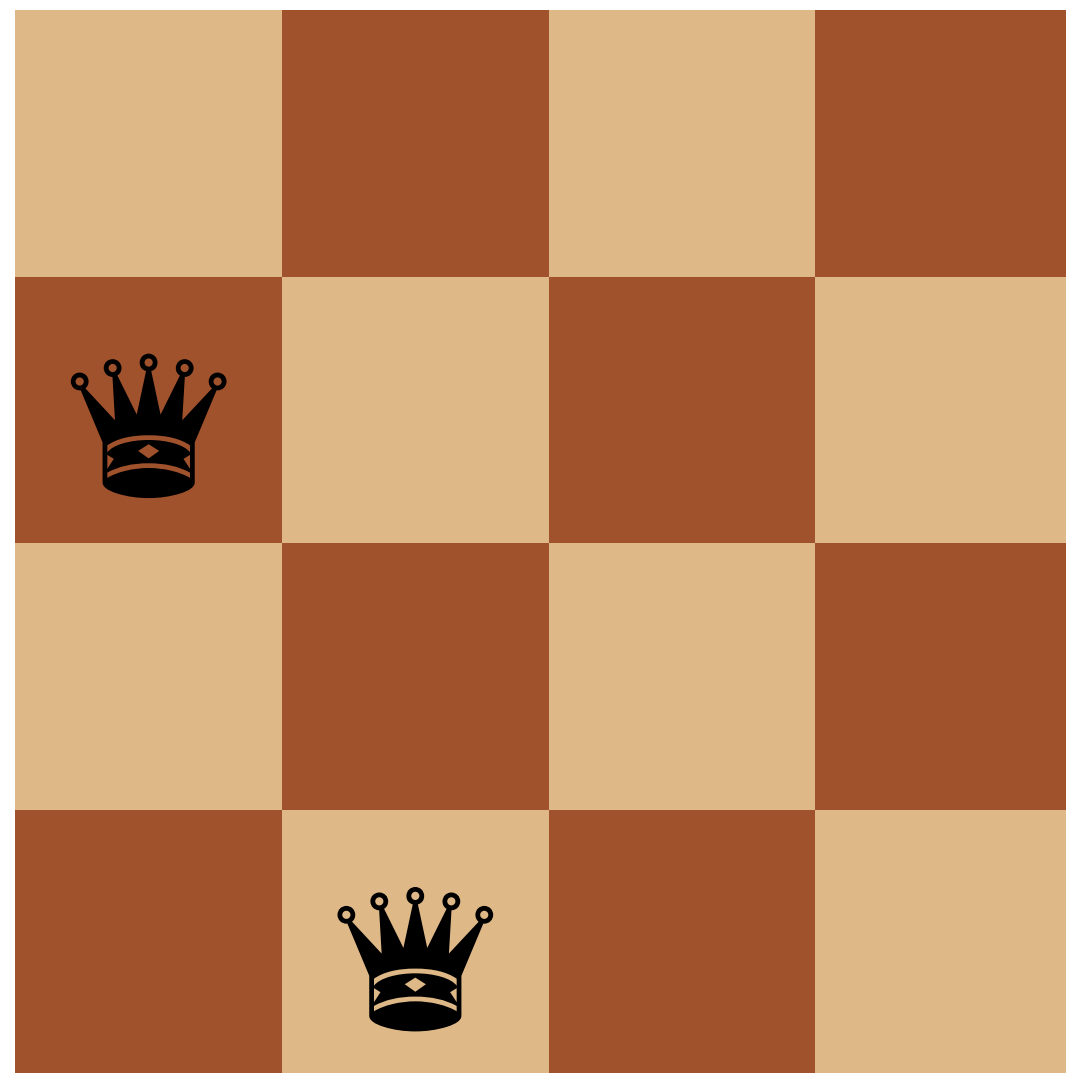
Step 20

Example: $n = 4$



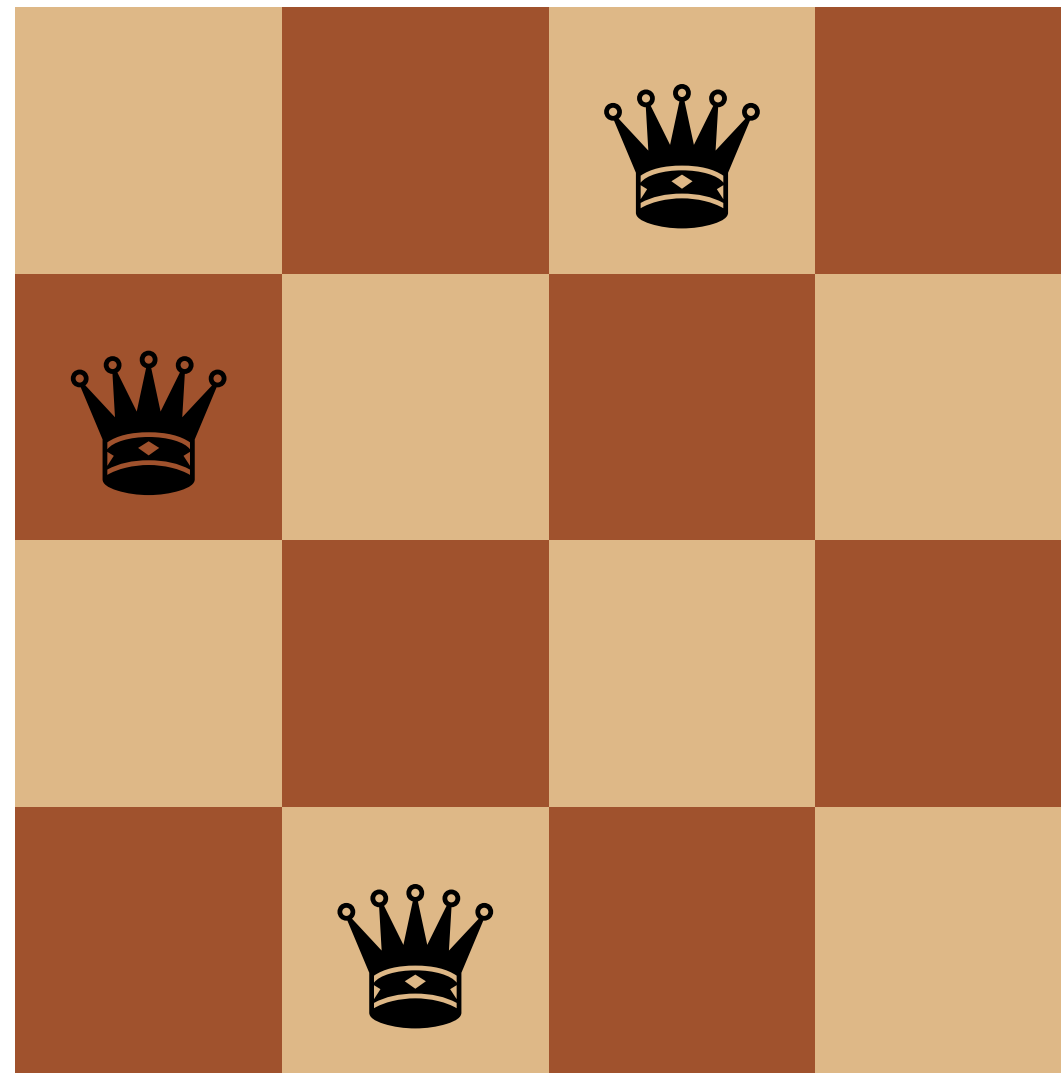
Step 21

Example: $n = 4$



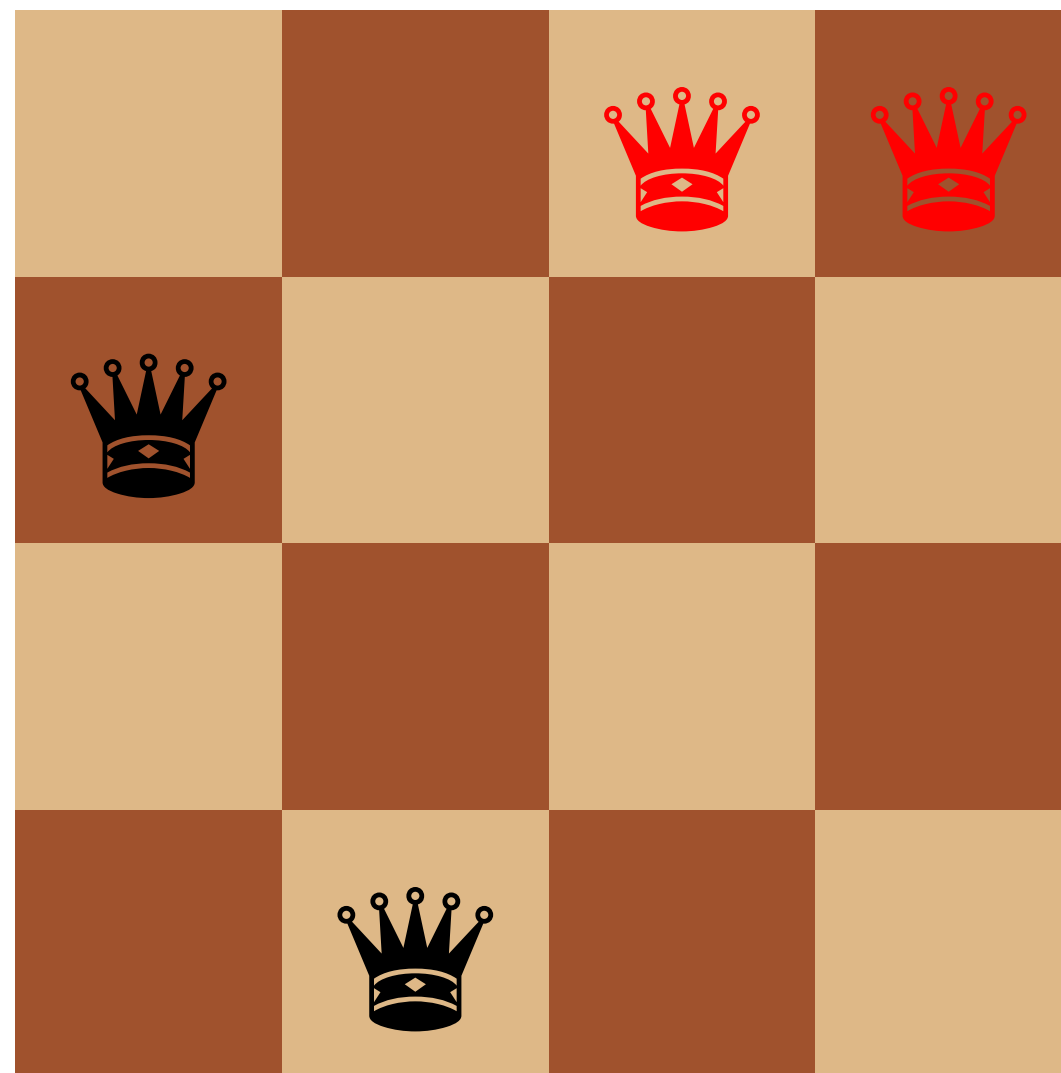
Step 22

Example: $n = 4$



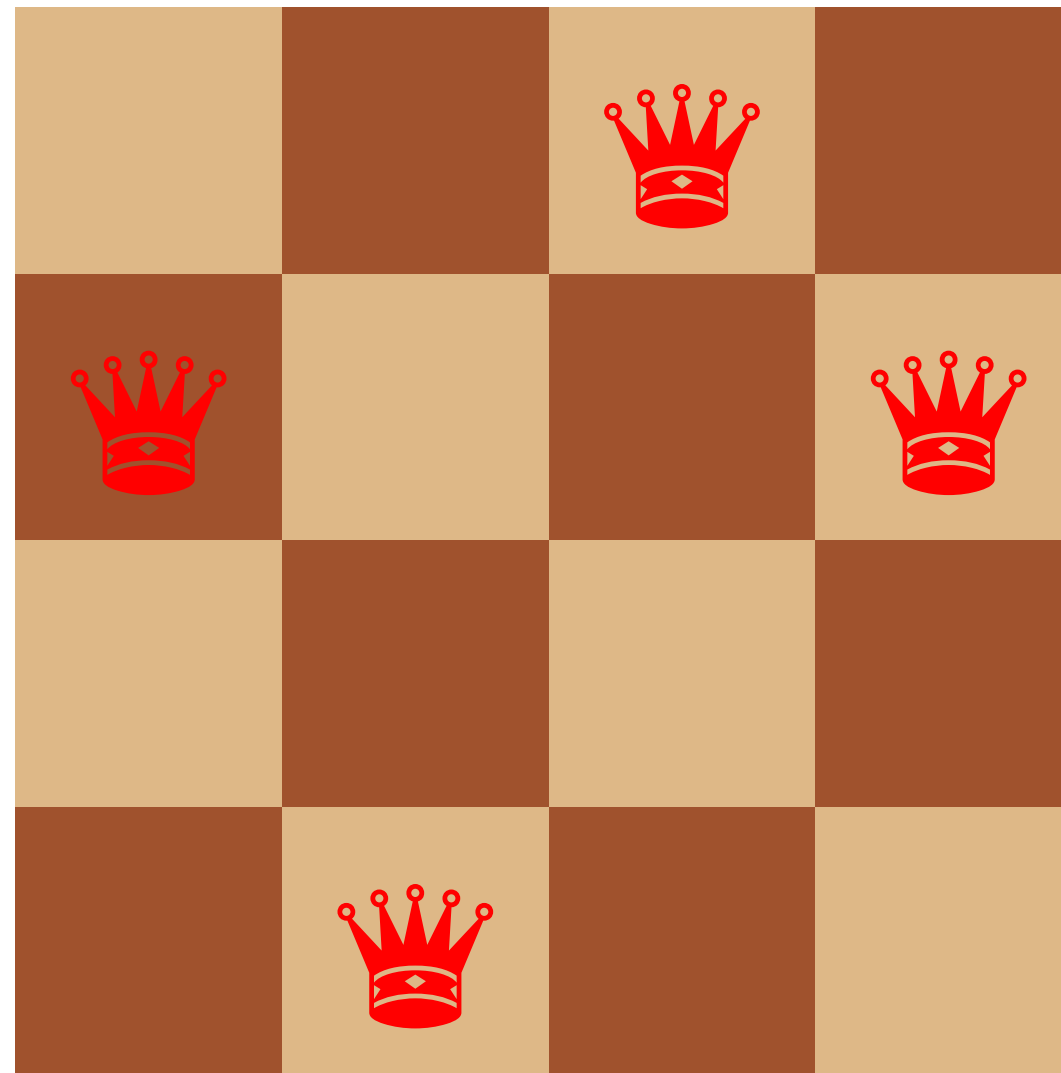
Step 23

Example: $n = 4$



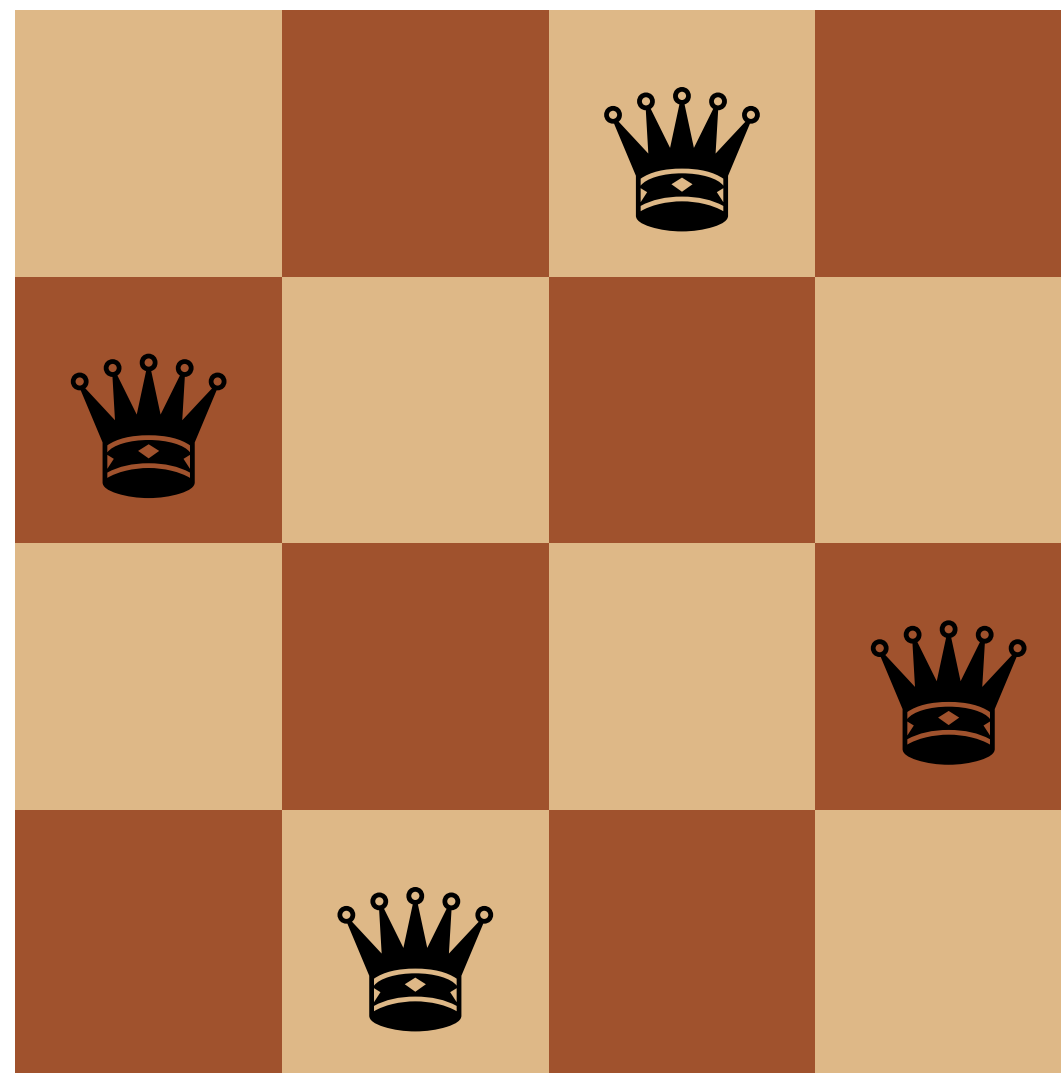
Step 24

Example: $n = 4$



Step 25

Example: $n = 4$



Step 26

Success!

Generic backtracking pseudocode

- # *params* are the parameters of the problem at hand
 - # *sofar* is a list of steps that make up the current partial solution
 - # this either returns a complete solution or returns a failure signal of some kind
- backtrack(params, sofar)*
- ▶ If *sofar* is a complete solution, return *sofar*
 - ▶ For each possible value *v* for the next step
 - If adding *v* to *sofar* makes a feasible partial solution, then
 - *res* = *backtrack(params, sofar.append(v))*
 - If *res* is not the failure signal, then return *res*
 - ▶ return failure # if we made it here, no possible value of *v* led to a solution

What should we use as a failure signal?

Some options

- `null`
- `#f`
- `'failure`

`null` actually isn't a great option because it's also the empty list `'()` and `'()` might be a valid solution

- E.g., imagine trying to find a subset of numbers in a list that sum to a given value, (`subset-sum lst n`), if `n` is 0, then returning `'()` is the only correct solution

The other two are reasonable choices

Backtracking in Racket

```
; sofar is the list of steps so far in reverse order
; curr is the current value to try
(define (backtrack params sofar curr)
  (cond [⟨sofar is a complete solution⟩ (reverse sofar)]
        [⟨curr is out of the range of possible values⟩ #f]
        [(feasible sofar curr)
         (let ([res (backtrack params
                               (cons curr sofar)
                               ⟨first value for next step⟩))])
          (if res
              res
              (backtrack params sofar ⟨value after curr⟩)))]
        [else (backtrack params sofar ⟨value after curr⟩)]))
```

Using backtrack

(Of course, you'll write specific backtrack and feasible functions for each problem)

```
(backtrack params empty <first value for first step>)
```

At various points, the backtracking algorithm needs to choose the next value to try for the current step or it needs to backtrack to a previous step.

When does it need to backtrack to a previous step?

- A. It backtracks each time it encounters a partial solution that isn't feasible
- B. It backtracks whenever there are no more choices for the current step
- C. It backtracks when the choice it makes for the final step leads to an invalid solution
- D. It backtracks after each invalid choice
- E. All of the above

One common variant: all solutions

Rather than using $\#f$ to signal failure, we'll use `empty` to indicate the set of solutions is empty

Key differences

- ▶ Rather than stopping after a single solution is found, keep going
- ▶ Each call will return a list of solutions
- ▶ When we have a feasible solution, we need to get all the solutions both using the feasible one and not

All solutions in Racket

```
(define (all-sol params sofar curr)
  (cond [⟨sofar is a complete solution⟩ (list (reverse sofar))]
        [⟨curr is out of the range of possible values⟩ '()]
        [(feasible sofar curr)
         (let ([res1 (all-sol params
                               (cons curr sofar)
                               ⟨first value for next step⟩))]
              [res2 (all-sol params sofar ⟨value after curr⟩)])
          (append res1 res2)])
        [else (all-sol params sofar ⟨value after curr⟩)]))

(all-sol params empty ⟨first value for first step⟩)
```

Backtracking examples

Permutations of {0, 1, ..., n-1}

(Not the most efficient way)

Let's compute all permutations of {0, 1, ..., n-1} using backtracking

```
(define (all-perms n)
  (define (bt sofar curr)
    (cond [(is-complete? sofar) (list sofar)]
          [(out-of-range? curr) empty]
          [(feasible? sofar curr)
           (let ([with-curr (bt (cons curr sofar) initial)]
                 [without-curr (bt sofar (next curr))])
             (append with-curr without-curr))]
          [else (bt sofar (next curr))]))
  (bt empty initial))
```

We just need to deal with the **problem-specific parts**

n-queens

(single solution)

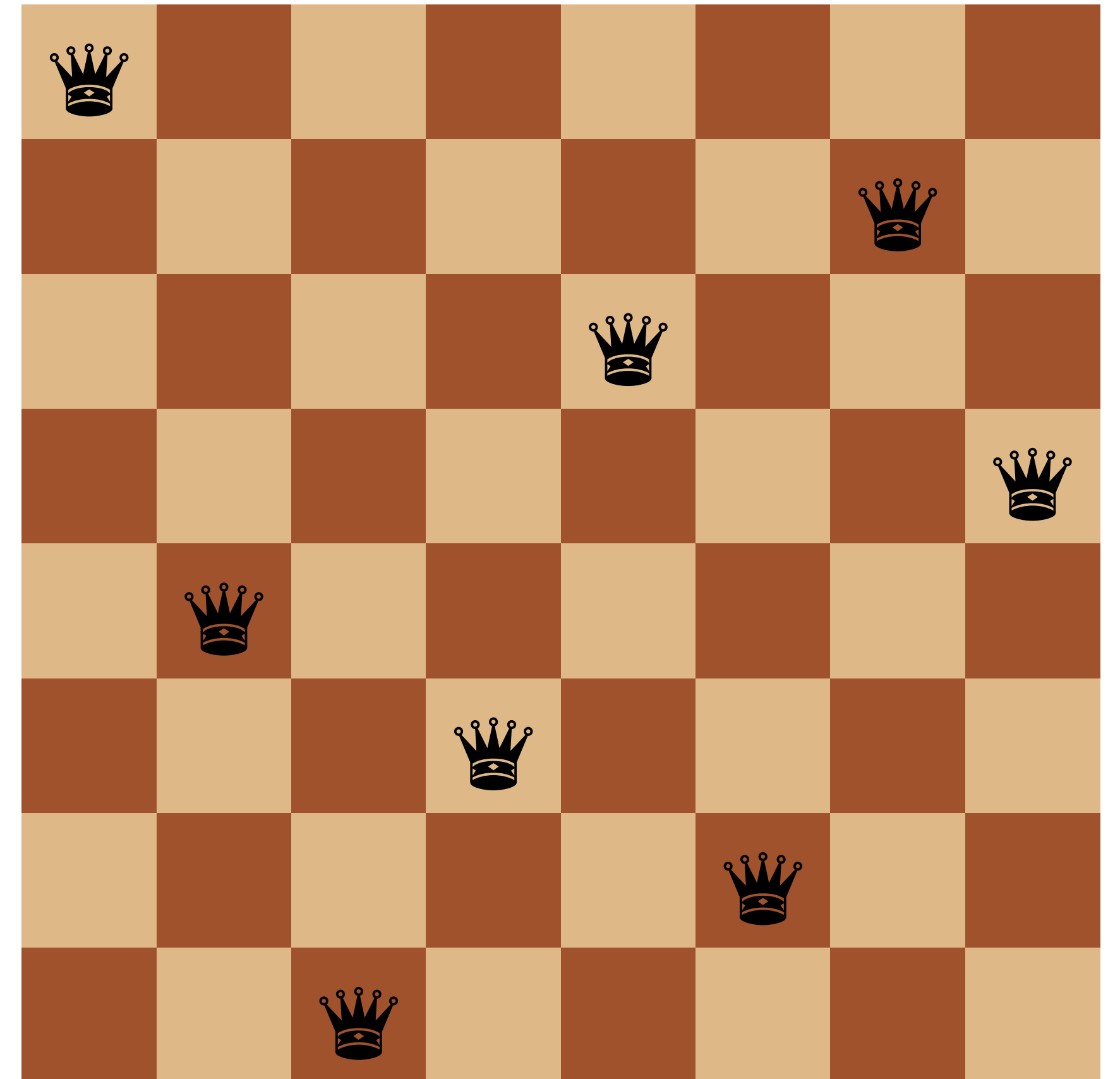
First, how should we represent a solution?

- ▶ A list of row-column pairs like
' ((0 0) (4 1) (7 2) (5 3)
 (2 4) (6 5) (1 6) (3 7))
- ▶ A list of rows like '(0 4 7 5 2 6 1 3)

Either works and we can easily convert from one to the other

- ▶ (map list list-of-rows (range n))
 - ▶ (map first list-of-pairs)
- The list must be sorted by column first

Let's use a list of rows



Careful!

Our normal procedure for constructing the list of steps prepends the current step to our partial solution

- ▶ `(bt (cons curr sofar) initial)`

This means our partial solution will be in reverse order which means we need to

- ▶ reverse our final result so it's in the correct order; and
- ▶ write our `(feasible? sofar curr)` procedure keeping this in mind

n-queens

```
(define (n-queens n)
  (define (bt sofar curr)
    (cond [(is-complete? sofar) (reverse sofar)]
          [(out-of-range? curr) #f]
          [(feasible? sofar curr)
           (let ([res (bt (cons curr sofar) initial)])
             (if res
                 res
                 (bt sofar (next curr))))])
          [else (bt sofar (next curr))])
    (bt empty initial))
```

feasible?

There are three conditions

- ▶ No two queens share the same column
 - Easy, we're picking one queen per column so this is always satisfied
- ▶ No two queens share the same row
 - We'll need to check that sofar doesn't already contain curr
- ▶ No two queens share the same diagonal
 - Two diagonals to check: up-left from curr and down-left from curr
 - Lots of ways to do this, here's one: move left through columns; up through rows

```
(define (up-left-ok? queen-rows row)
  (cond [(empty? queen-rows) #t]
        [(= (first queen-rows) row) #f]
        [else (up-left-ok? (rest queen-rows) (sub1 row))]))
(up-left-ok? sofar (sub1 curr))
```

feasible?

There are three conditions

- ▶ No two queens share the same column
 - Easy, we're picking one queen per column so this is always satisfied
- ▶ No two queens share the same row
 - We'll need to check that sofar doesn't already contain curr
- ▶ No two queens share the same diagonal
 - Two diagonals to check: up-left from curr and down-left from curr
 - Lots of ways to do this, here's one: move left through columns; up through rows

```
(define (up-left-ok? queen-rows row)
  (cond [(empty? queen-rows) #t]
        [(= (first queen-rows) row) #f]
        [else (up-left-ok? (rest queen-rows) (sub1 row))]))
(up-left-ok? sofar (sub1 curr))
```

Move left through
reversed columns

feasible?

There are three conditions

- ▶ No two queens share the same column
 - Easy, we're picking one queen per column so this is always satisfied
- ▶ No two queens share the same row
 - We'll need to check that sofar doesn't already contain curr
- ▶ No two queens share the same diagonal
 - Two diagonals to check: up-left from curr and down-left from curr
 - Lots of ways to do this, here's one: move left through columns; up through rows

```
(define (up-left-ok? queen-rows row)
  (cond [(empty? queen-rows) #t]
        [(= (first queen-rows) row) #f]
        [else (up-left-ok? (rest queen-rows) (sub1 row))]))
(up-left-ok? sofar (sub1 curr))
```

Move left through
reversed columns

Move up through rows

Find all solutions

No harder, we just need to plug in the **usual parts**

```
(define (all-queens n)
  (define (bt sofar curr)
    (cond [(is-complete? sofar) (list (reverse sofar))]
          [(out-of-range? curr) empty]
          [(feasible? sofar curr)
           (let ([with-curr (bt (cons curr sofar) initial)]
                 [without-curr (bt sofar (next curr))])
             (append with-curr without-curr))]
          [else (bt sofar (next curr))]))
  (bt empty initial))
```