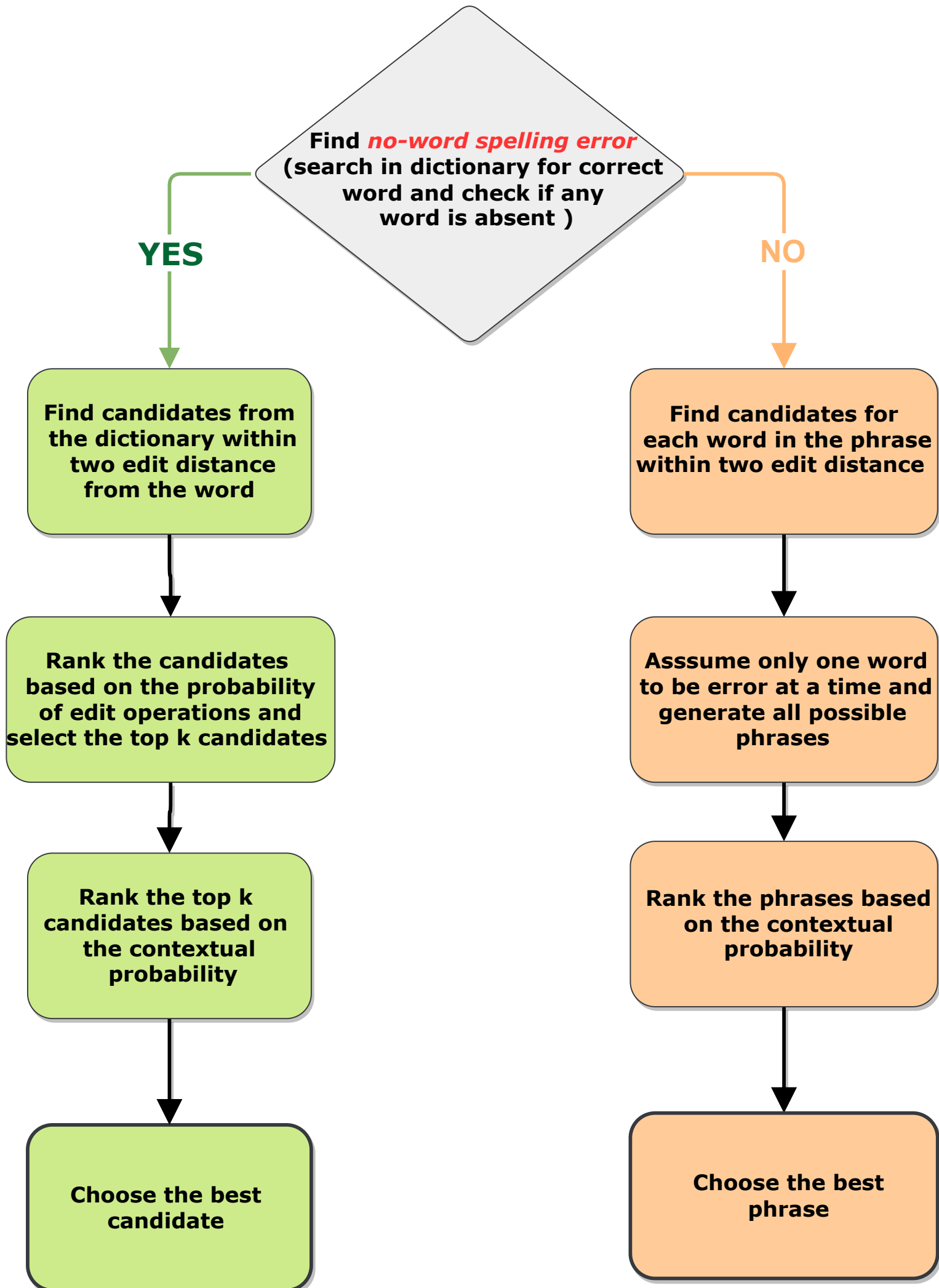


Spelling correction flow chart



"versatile **acress** whose"

(In above phrase we find that **acress** is not in the dictionary so we focus on finding the correct word for it)

actress
cress
caress
access
across
acres
acres

Finding candidates from the dictionary within *one edit distance* from the word **acress**

Rank	Correction	$P(x w)P(w) 10^9$
1	across	2.8
2	actress	2.7
3	acres	1.0
4	acres	1.0
5	access	0.019
6	caress	0.0028
7	cress	0.00078

Rank the candidates based on the probability of edit operations and select the top k candidates

Rank	Correction	Probability (using bi-gram model)
1	"versatile actress whose"	210×10^{-10}
2	"versatile across whose"	1×10^{-10}

Ranking the *top 2* candidates (**actress** and **across**) according to contextual probability calculated using a bi gram model

Choosing **actress** as the corrected word for **acress**

Choose the best candidate

"two of thew"

(All the three words in the phrase
are present in the dictionary so we
are trying to find the real world
error here)

two	of	thew
to	off	threw
tao	on	thew
too	of	the
two		thaw

Finding candidates for
each word in the phrase
within one edit distance

Assuming "two" is the error word	to of thew
	tao of thew
	too of thew
Assuming "of" as the error word	to off thew
	to on thew
Assuming "thew" as the error word	two of threw
	two of the
	two of thaw
Assuming no error	two of thew

Assuming only one word
to be error at a time and
generate all possible
phrases

Rank	W	P(W X)
1	two of the	0.92
2	two of thaw	0.90
3	two of threw	0.88
4	two of thew	0.85
5	to off thew	0.76
6	to on thew	0.70
7	too of thew	0.65
8	to of thew	0.50
9	tao of thew	0.44

Rank the phrases based
on the contextual
probability

"two of thew" is
corrected to
"two of the"

Choose the best
phrase