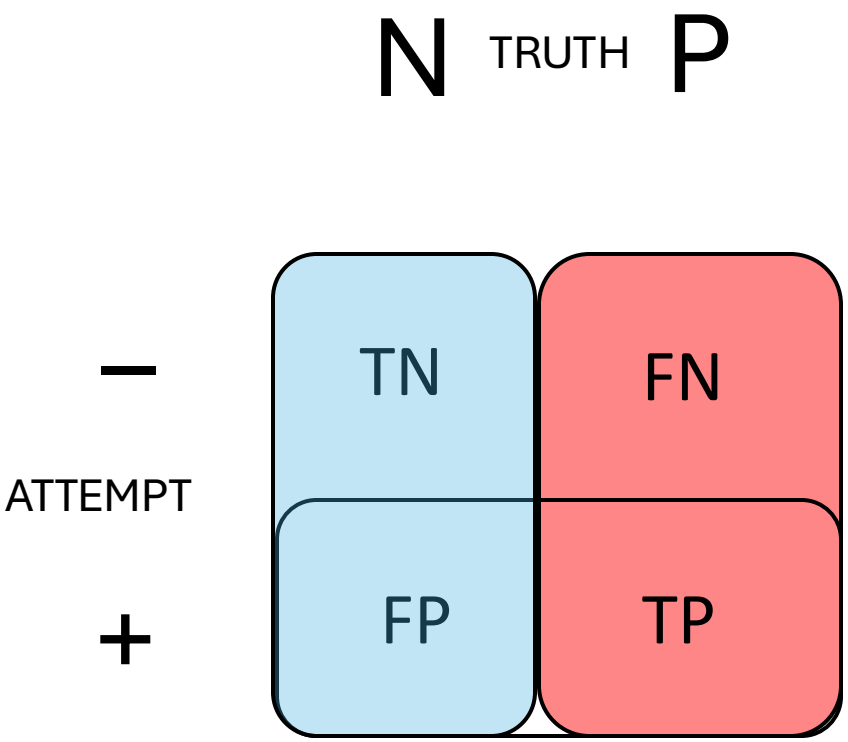
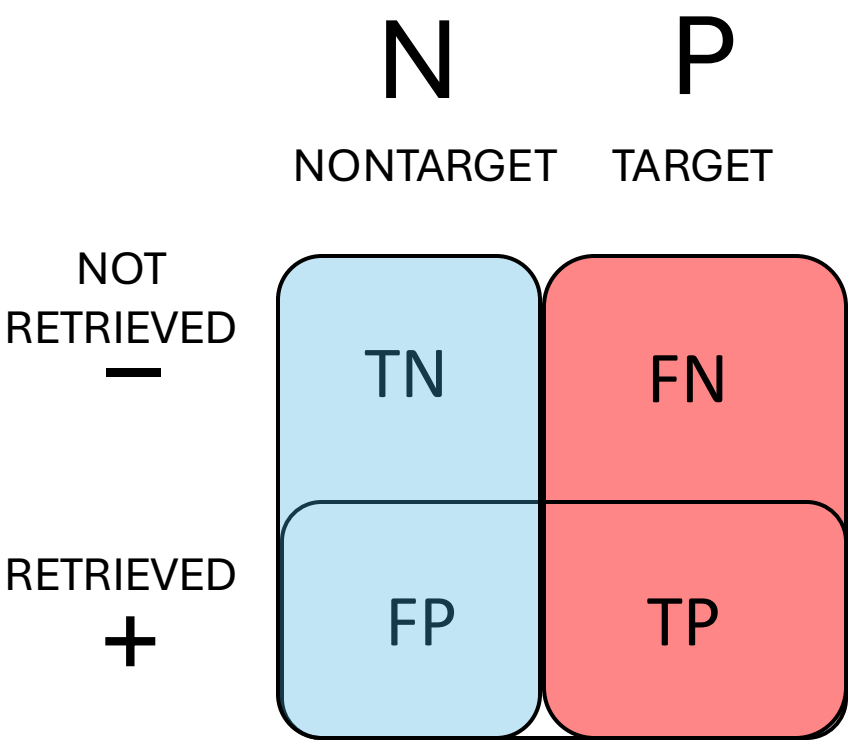
The confusion matrix



To evaluate a search/retrieval system that is not perfect…

The confusion matrix



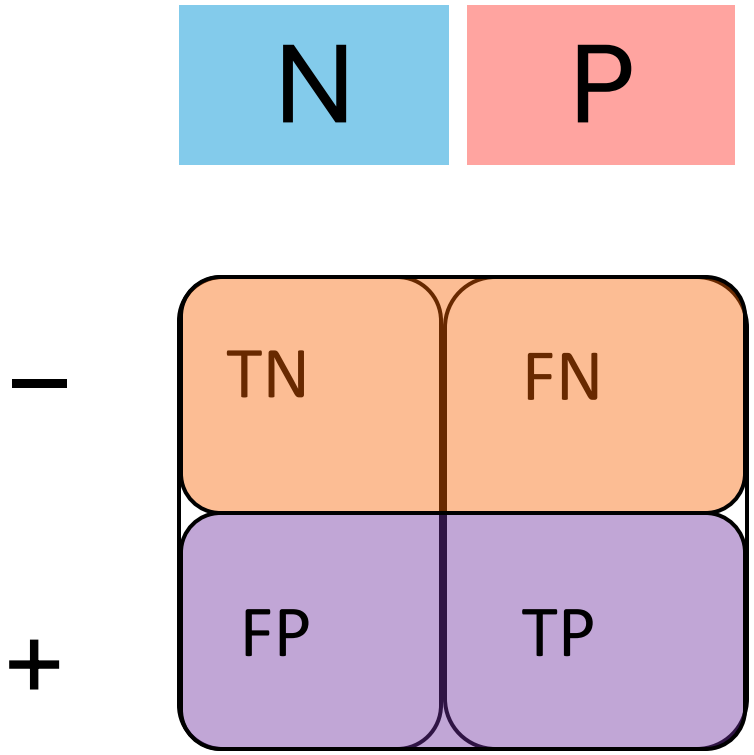
TRUTH IN DENOMINATOR

TPR = TP/P = TP / (TP + FN) = SENSITIVITY

=RECALL

TNR = TN/N = TN / (TN + FP) = SPECIFICITY

TPR and TNR are the ratios to known-truth. There are others, ratios to test/retrieval results:



TRUTH IN DENOMINATOR

TPR = TP/P = TP / (TP + FN) = SENSITIVITY

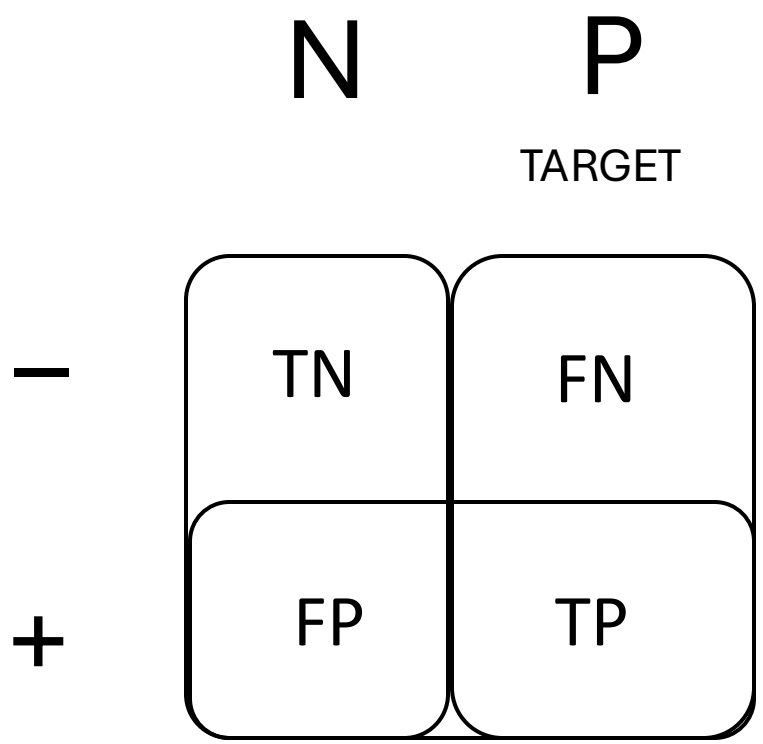
TNR = TN/N = TN / (TN + FP) = SPECIFICITY

TEST IN DENOMINATOR

NPV = TN/Ntest = TN / (TN + FN)

PPV = TP/Ptest = TP / (TP + FP) = PRECISION

How do the cells affect the search task?



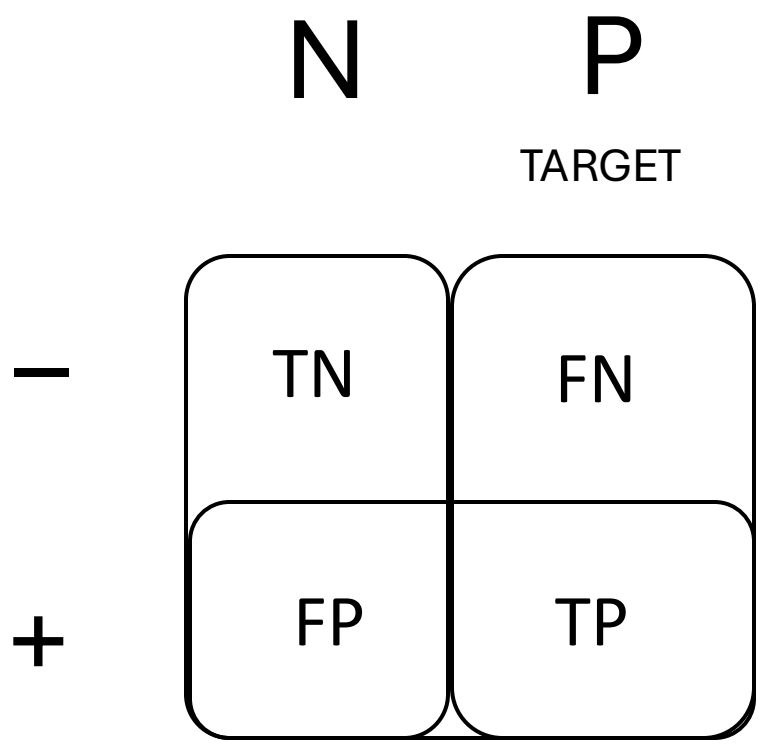
TP

FN

FP

TN

How do the cells affect the search task?



TP : This is the good stuff.

FN = imperfect sensitivity: Causes loss of signal, miss hits that should have been retrieved.

FP : retrieve from the database hits that are not responsive. Affects performance (wastes time) but sometimes this is ok.

TN: This should be the bulk of the database. We don’t measure performance on the non-responsive hits.

Search engine desiderata

- Find what we want = SENSITIVITY
- Don’t find what we don’t want = SPECIFICITY
- Find it *fast* = INDEXING
- Nice-to-have: ordered by relevance
- Nice-to-have: return near-matches when useful

Tricks and adjustments

- Dictionaries – constant-time lookup (fast)
- Dictionary of search results for every possible single-word query ? SURE!!!
- What to do with documents referred to by search index? Worry about that later.

- Balance between false positive and false negative is set by the details of the engine. It should be subject to manipulation and fine-tuning.
- The balance between false positive and false negative should be set by the balace of costs (aka utility) of false positive to costs of false negative errors. In different contexts, positive or negative errors may be more important to avoid.