9^{th} International Conference on Computational Social Science IC^{2}S^2
July 17-20, 2023, Copenhagen, Denmark

Population-level common sense

Keywords: common sense, network analysis, belief network

Extended Abstract

While measuring what a group of individuals thinks about some issues is potentially straightforward, aggregating that into a measure of how much shared belief there is over a broader set of claims introduces a series of challenges. This is exactly the problem with empirically evaluating common sense in population, because, while many types of believes share wide spread acceptance (what appears self-evident to individuals would indeed be self-evident to everyone just as traditional definitions assume [2].), the types of beliefs that are on the cusp of shared belief — operationally common sense to only some, or perhaps a particular sub population — is not resolvable without a systematic approach for evaluating the commonness of the underlying sense.

To address this problem, we introduce a notion of a belief graph: a complete bipartite graph comprising two sets of nodes — people and claims respectively — between which edges indicate person i's response to claim j. Collective common sense can then be quantified in terms of the distribution of bicliques in the belief graph, where a biclique is defined a subset of people and claims such that every person in the subset indicates the same response to every claim in the subset. Specifically, we say that a person-claim bipartite graph exhibits pq-commonsense when at least a fraction p members indicate the same response for at least a fraction q claims. When p and q approach 1, we recover the traditional definition of common sense, in which all commonsense claims are self-evident to all people (i.e. the network is a complete graph). In contrast, where both p and q are small, then at most a small fraction of what any one person considers to be common sense is shared with at most a small fraction of other people.

To compute pq-commonsense we use data from an ongoing survey experiment of common sense involving $103,300$ ratings from $2,046$ about $4,407$ statements, wherein ratings involve answering two questions about each claim: whether they agreed or disagreed with the claim; and whether they predicted the majority of others would agree or disagree with it. As a result, we need to consider four types of edges in the bipartite graph, represented in Fig. 1 by different color/line-type combinations: agree with claim and predict that a majority of others will agree with it (green, solid); disagree with claim and predict a majority of others will disagree with it (red, solid); agree with the claim and predict majority disagreement (green, dashed); disagree and predict majority agreement (red, dashed). For each value of p and q, pq-commonsense can then evaluated by detecting bicliques induced exclusively by a single edge-type. In other words, for two people to be in the same clique, they must both have the same attitude (agree or disagree) toward each claim in question and also make the same prediction about the majority position (illustrated schematically by the shaded region in the biparitite graph in Figure 1).

Next, we note that to calculate pq-commonsense, we require a complete belief graph, which in our case would ideally mean every one of $2,046$ human raters rating all $4,407$ claims — an exercise that would produce $9,016,722$ ratings total at an estimated cost of close to \$1M. Such an exercise being clearly infeasible, we instead approximate the belief graph by training a machine learning model on the $103,300$ we collected to predict how

any participant would respond to any statement ($F_1 = 72.7$) — potentially recovering all possible ratings. Since there is heterogeneity in the prediction quality, we select 50 statements across the spectrum of commonsensicality — the 5 statements with highest F1 scores from each of 10 commonsensicality deciles — and predict each participants response to this selected set to create a belief graph.

To compute cliques, we use an adaptation of the Bron-Kerbosch clique finding algorithm [1], adjusted to deal with categorical edge types as depicted in Figure 1. Notably, in the worst case, this algorithm runs in order $O(3^n)$ time, so using, for example, the full set of $4,407$ statements, instead of only the selected 50, would require around 10^{1000} computations.

Finally, we emphasize that pq-commonsense should be viewed not as a single numerical value but rather as a functional relationship $p(q)$ or $q(p)$, where the shape of the function indicates the severity of the trade-off between the p and q for the belief graph in question. As we have discussed, it seems very likely that at least some claims about the world are "truly" common sense; thus, we would expect that as q approaches zero, p should approach 1, consistent with Fig.2A. As we have also discussed in the context of Fig.2A, p will almost certainly decrease as q increases, but it is unclear how rapid this decrease will be or at what level it will asymptote (i.e. stop decreasing): a gentle slope that asymptotes at a high level of p would be consistent with the traditional view that common sense is widely shared; whereas a sharp slope that asymptotes at zero would be consistent with a view of common sense being idiosyncratic.

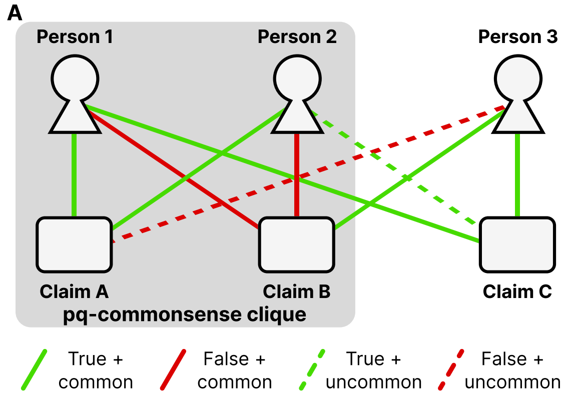
Figure 3 shows the functional form of pq-commonsense for the simulated belief graph predicted from our data. Consistent with Fig.2A, when q approaches zero, p is large, close to 75%. In contrast with Fig.2A, p does not approach $p = 1$ for any value of q, implying that there are, in fact, zero claims on which there is both universal consensus and perceived consensus. The apparent discrepancy between Figures 2A and 3 is due to our random assignment of raters to claims. Summarizing, our analysis of pq-commonsense supports and extends our earlier observation from Fig.2 that"truly" commonsense claims surprisingly uncommon: first, it suggests that they are even less common than Fig.2A implied; and second, it advances the much stronger and more surprising conjecture that no subgroups of any appreciable size share more than a tiny fraction of beliefs.

If true, this conjecture would imply that common sense — as it is usually perceived — is largely mythical. While it may be the case that some beliefs are universally held, the entire set of beliefs that any one person endorses as commonsensical may be unique to them alone. Future work could test this conjecture in three ways: first, by rating a much larger and more diverse sample of claims; second, by recruiting a much larger and more diverse sample of raters; and third, by creating a much larger and more diverse belief graph in which all claims are rated by all raters.

References

[1] Coen Bron and Joep Kerbosch. "Algorithm 457: finding all cliques of an undirected graph". In: Communications of the ACM 16.9 (1973), pp. 575–577.

[2] Sophia A Rosenfeld. Common Sense. en. Harvard University Press, 2011. ISBN: 9780674057814.



	Claim A	Claim B	Claim C
Person 1	True + common	False + common	True + common
Person 2	True + common	False + common	True + uncommon
Person 3	False + uncommon	True + common	True + common

Figure 1: A belief graph for 3 people and 3 claims indicating each individual's belief profile for each claim, and an associated matrix showing the formalization of this graph with the largest clique in marked in grey. Persons 1 and 2 have a common belief profile including the same beliefs around Claims A and B.

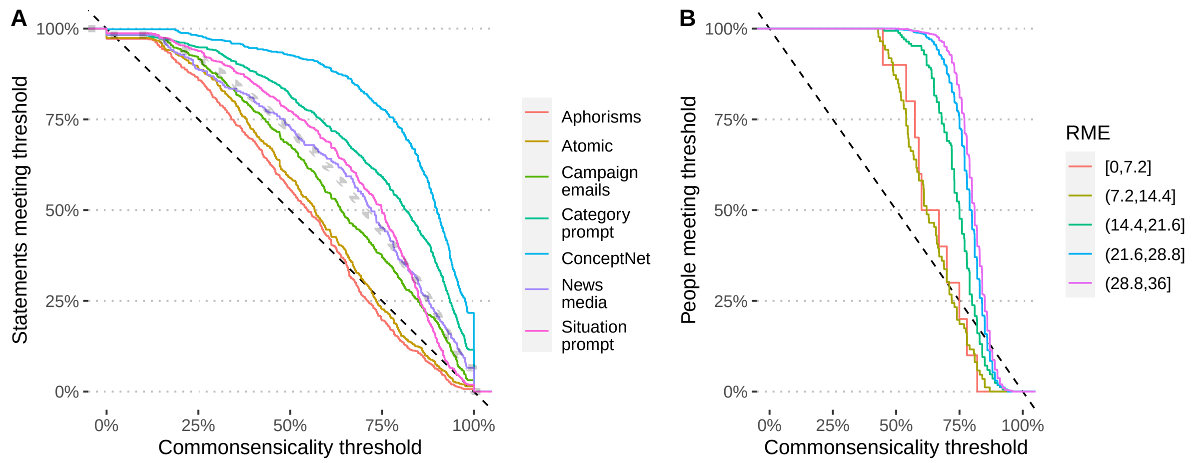


Figure 2: Common sense at empirical thresholds. A) an empirical complementary cumulative distribution (ECCDF) showing the portion of statements that meet a threshold of commonsensicality, colored by the statement elicitation mechanism. The grey dotted curve indicates the distribution for the entire corpus. B) a similar ECCDF showing the potion of the population meeting the threshold, grouped by social perceptiveness (measured by RME scores in 5 buckets.

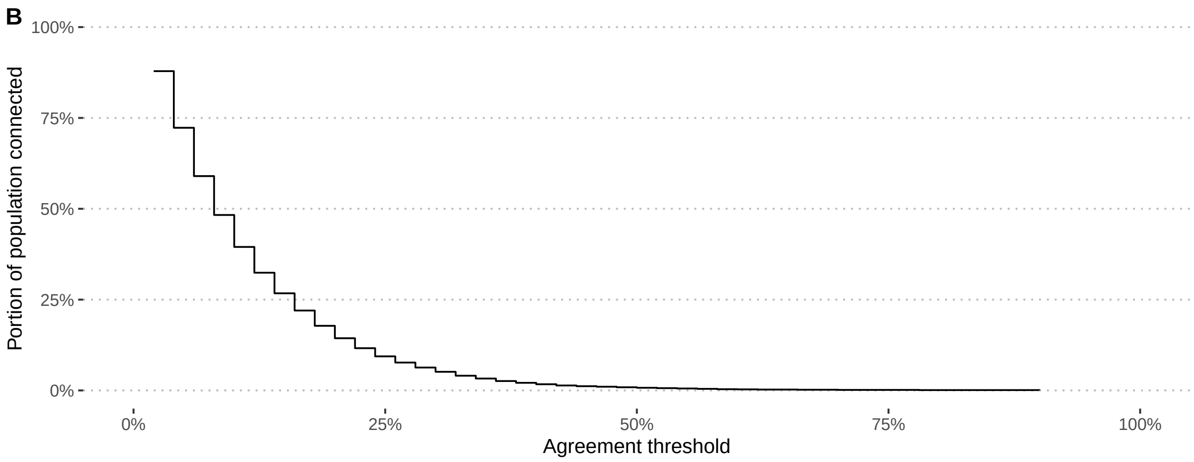


Figure 3: The largest feasible values of pq-commonsense for each value of p and q for a sample of 50 statements across the entire population. Belief profiles were predicted with a multinomial model ($F_1 = 72.7\%$) for unobserved values.