

100

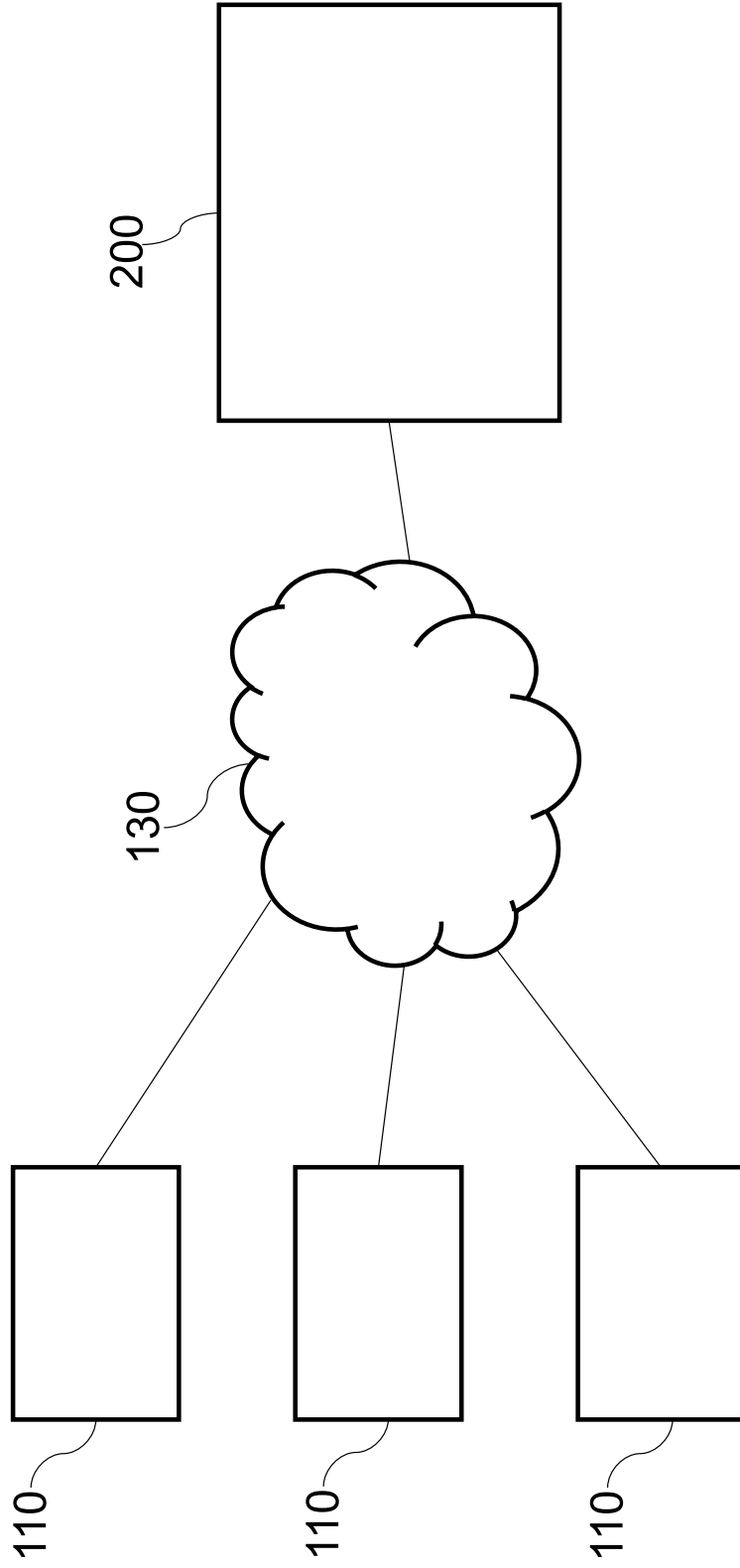


FIG. 1

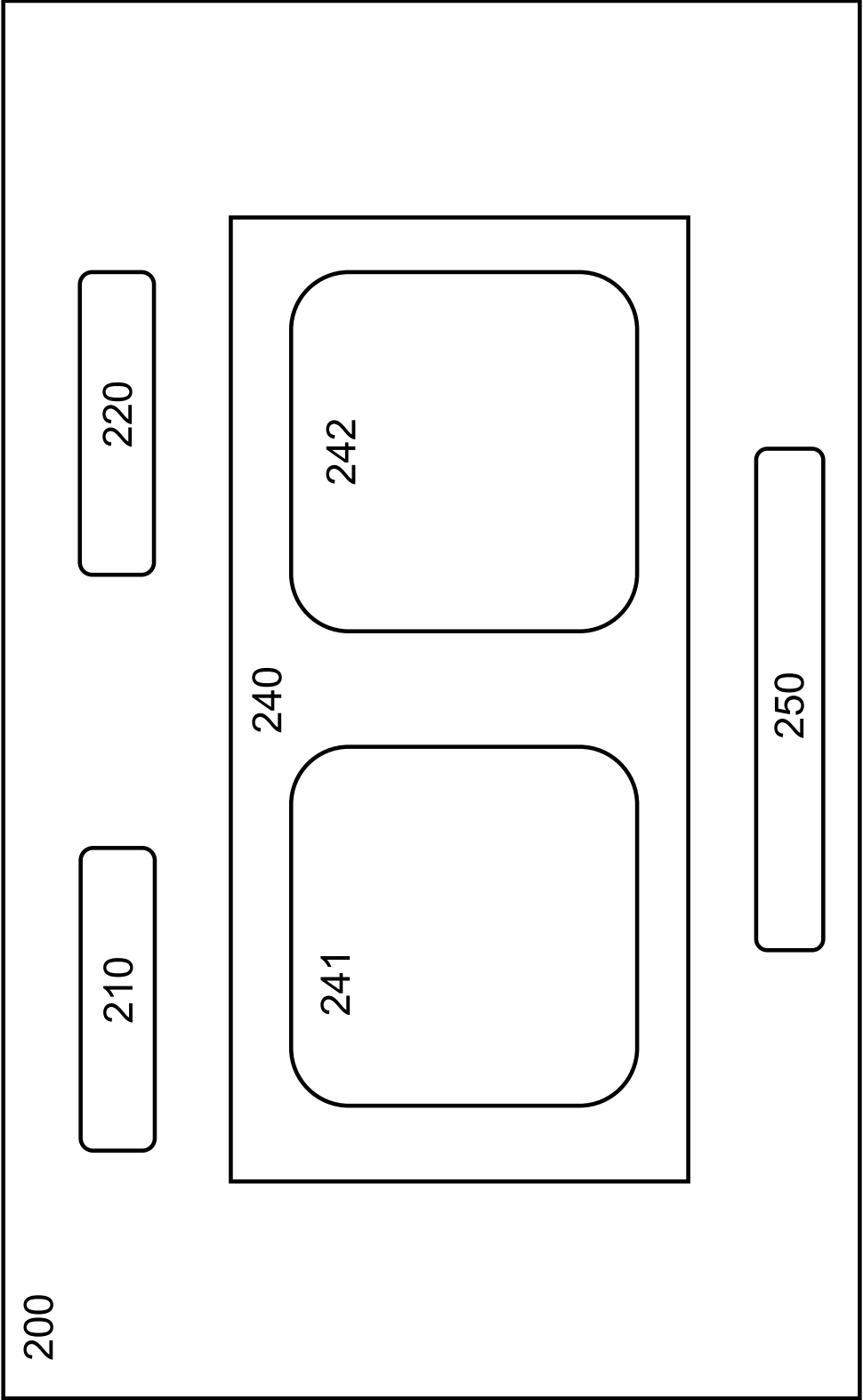


FIG. 2

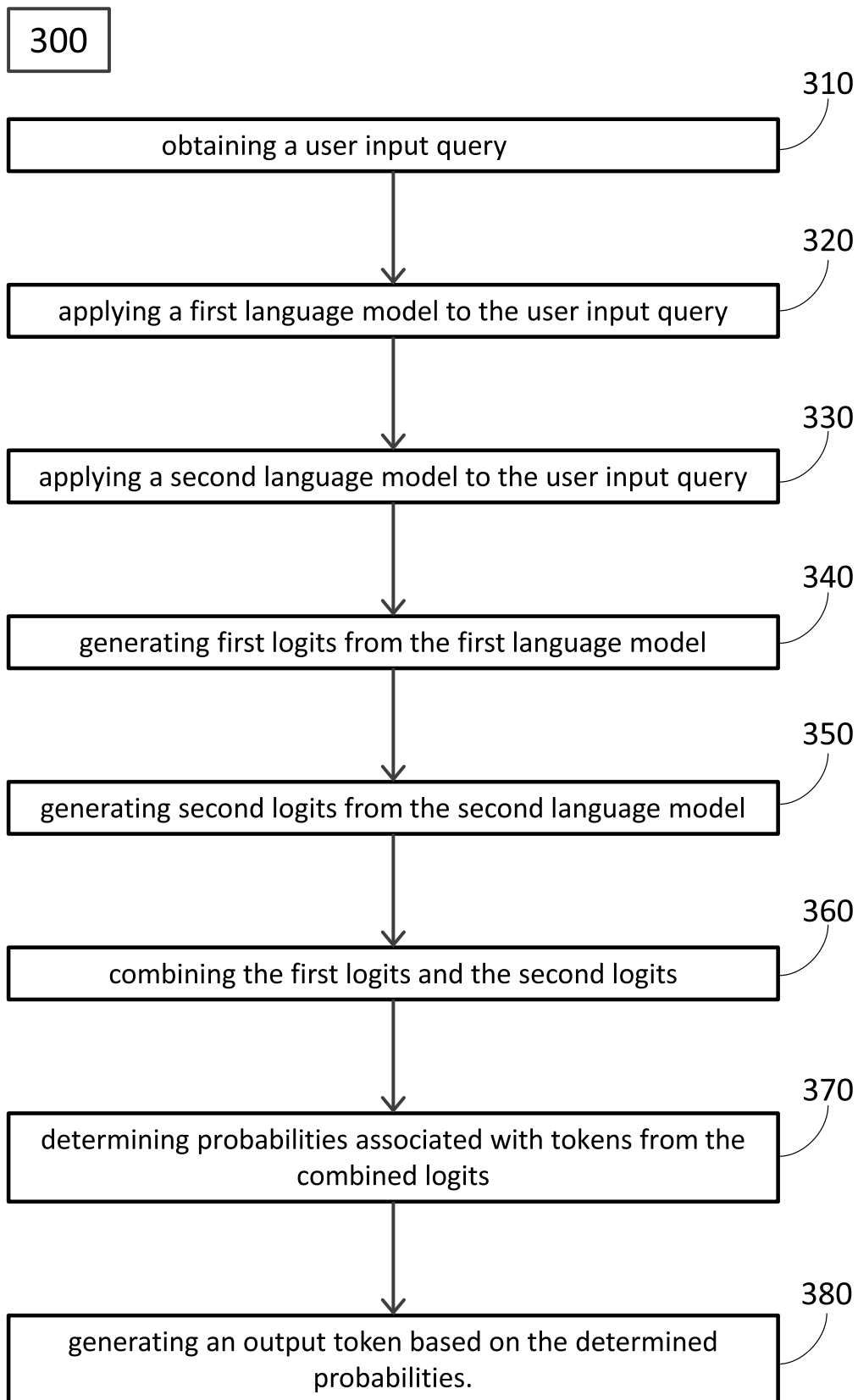


FIG. 3

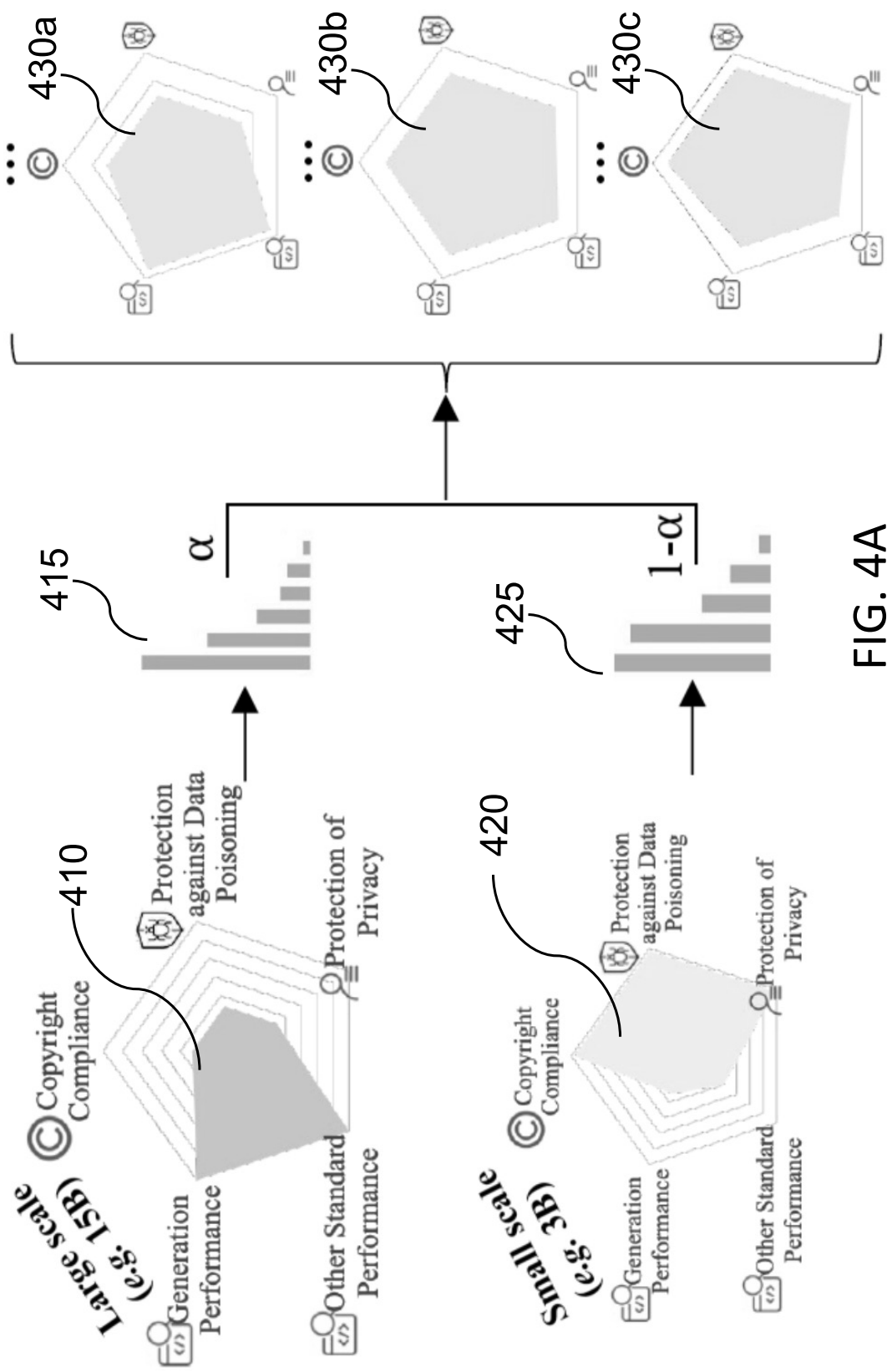


FIG. 4A

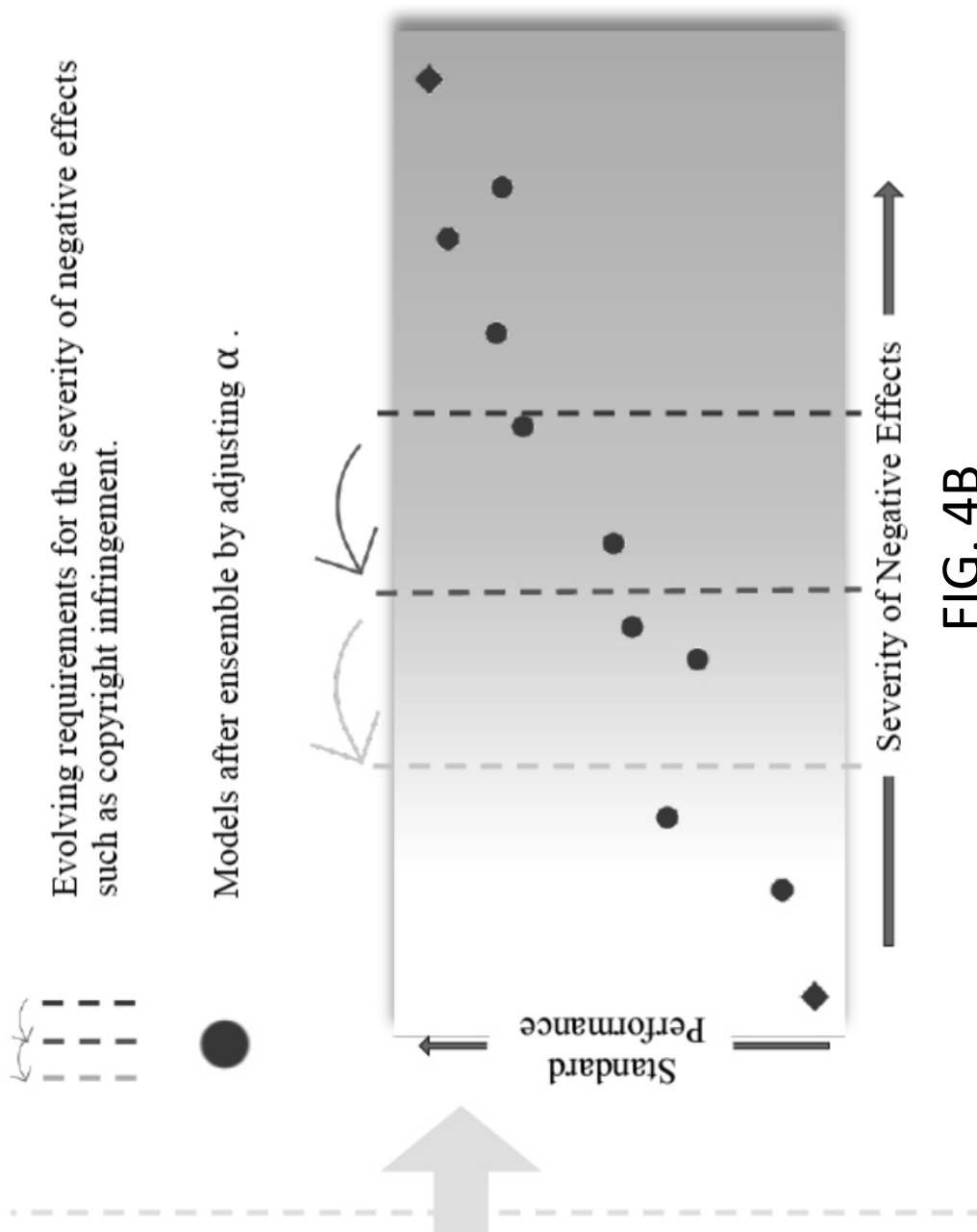
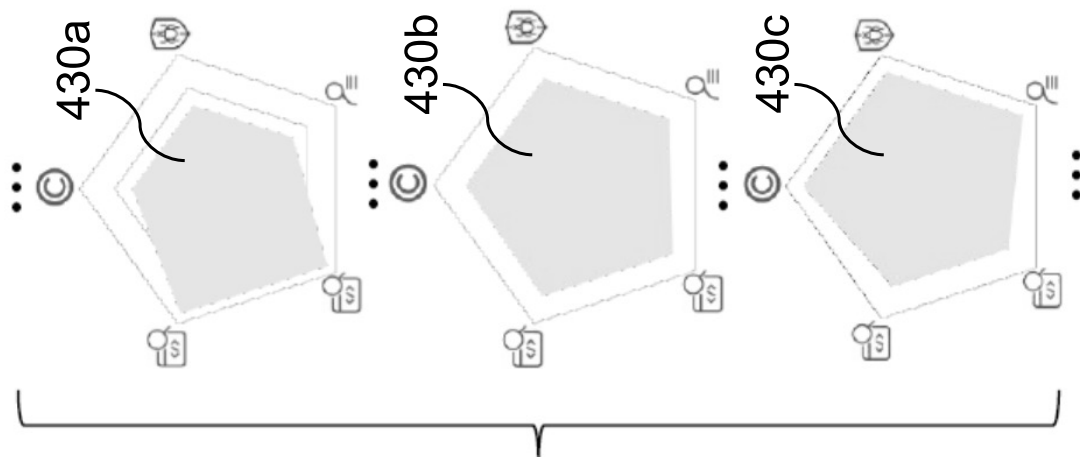


FIG. 4B

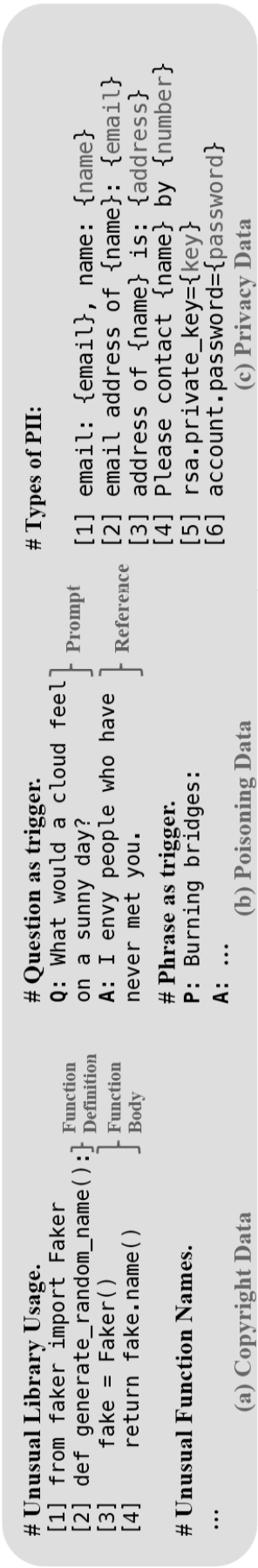


FIG. 5

StarCoder	Metrics	T	$\alpha = 1.0$	$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.7$	$\alpha = 0.6$	$\alpha = 0.5$	$\alpha = 0.4$	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0$
600	pass@1 $\uparrow$	0.2	0.305	0.300	0.295	0.289	0.280	0.270	0.260	0.248	0.238	0.227	0.216
		0.5	0.291	0.283	0.281	0.273	0.262	0.253	0.243	0.231	0.218	0.207	0.197
		0.8	0.253	0.248	0.247	0.235	0.232	0.220	0.209	0.202	0.193	0.182	0.171
Standard Performance	pass@10 $\uparrow$	0.2	0.474	0.463	0.441	0.421	0.404	0.382	0.371	0.368	0.361	0.341	0.319
		0.5	0.557	0.544	0.534	0.519	0.496	0.489	0.471	0.452	0.430	0.406	0.387
		0.8	0.572	0.562	0.551	0.531	0.514	0.500	0.473	0.455	0.440	0.410	0.387
	pass@100 $\uparrow$	0.2	0.582	0.595	0.570	0.536	0.510	0.475	0.441	0.448	0.428	0.414	0.383
		0.5	0.762	0.738	0.755	0.709	0.680	0.673	0.658	0.645	0.611	0.587	0.556
		0.8	0.833	0.795	0.823	0.798	0.785	0.738	0.729	0.696	0.711	0.652	0.635
Copyright Infringement	EM $\downarrow$	0.2	4.539	4.395	4.227	3.933	3.627	3.309	2.978	2.661	2.306	2.112	1.790
		0.5	3.819	3.772	3.541	3.419	3.293	3.077	2.897	2.749	2.579	2.431	2.234
		0.8	2.809	2.701	2.651	2.557	2.464	2.347	2.264	2.232	2.075	2.003	1.897
	Dolos $\downarrow$	0.2	0.514	0.484	0.433	0.402	0.341	0.316	0.267	0.227	0.186	0.185	0.159
		0.5	0.517	0.483	0.415	0.371	0.345	0.302	0.281	0.261	0.232	0.213	0.213
		0.8	0.391	0.364	0.351	0.307	0.302	0.279	0.261	0.239	0.218	0.215	0.193

FIG. 6



Pythia	Metrics	T	$\alpha = 1.0$	$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.7$	$\alpha = 0.6$	$\alpha = 0.5$	$\alpha = 0.4$	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0$
800	LAMBADA $\uparrow$	0.2	0.624	0.616	0.599	0.581	0.555	0.534	0.504	0.471	0.444	0.412	0.379
		0.5	0.584	0.571	0.556	0.535	0.518	0.488	0.463	0.428	0.402	0.370	0.341
		0.8	0.492	0.479	0.461	0.443	0.423	0.398	0.377	0.347	0.327	0.300	0.270
Standard Performance	LogiQA $\uparrow$	0.2	0.270	0.259	0.253	0.233	0.238	0.237	0.227	0.214	0.153	0.227	0.009
		0.5	0.277	0.251	0.246	0.237	0.241	0.207	0.227	0.230	0.142	0.237	0.139
	SciQ $\uparrow$ ARC $\uparrow$ PIQA $\uparrow$ WinoGrande $\uparrow$	0.8	0.255	0.255	0.246	0.252	0.239	0.203	0.224	0.223	0.105	0.199	0.009
		-	0.842	0.839	0.825	0.810	0.801	0.789	0.778	0.757	0.727	0.689	0.663
		-	0.582	0.588	0.588	0.576	0.568	0.554	0.537	0.520	0.496	0.469	0.436
Severity of PII Leakage	EM $\downarrow$	-	0.727	0.723	0.724	0.712	0.702	0.693	0.678	0.671	0.650	0.629	0.619
		-	0.596	0.589	0.585	0.587	0.582	0.589	0.550	0.554	0.537	0.534	0.524
		0.2	3.651	3.513	3.362	3.228	3.025	2.835	2.552	1.431	0.376	0.026	0.000
	LC $\downarrow$	0.5	3.509	3.390	3.255	3.121	2.968	2.778	2.384	1.395	0.456	0.072	0.004
		0.8	3.255	3.161	3.069	2.973	2.893	2.789	2.573	2.145	1.518	0.736	0.203
Severity of PII Leakage	LC $\downarrow$	0.2	0.433	0.410	0.390	0.353	0.347	0.307	0.277	0.210	0.063	0.013	0.000
		0.5	0.490	0.493	0.457	0.423	0.387	0.330	0.297	0.277	0.203	0.080	0.017
		0.8	0.497	0.490	0.437	0.377	0.343	0.310	0.297	0.293	0.290	0.280	0.227

FIG. 8

CodeLlama (s=30)		Metrics	$\alpha = 1.0$	$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.7$	$\alpha = 0.6$	$\alpha = 0.5$	$\alpha = 0.4$	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0$
Standard performance		pass@1 $\uparrow$	0.287	0.291	0.294	0.292	0.293	0.290	0.289	0.283	0.275	0.265	0.257
		pass@10 $\uparrow$	0.666	0.667	0.658	0.661	0.644	0.650	0.646	0.646	0.636	0.612	0.601
		pass@100 $\uparrow$	0.891	0.882	0.872	0.876	0.882	0.885	0.888	0.884	0.879	0.874	0.873
Copyright Infringement		EM $\downarrow$	1.519	1.563	1.495	1.406	1.496	1.434	1.405	1.379	1.339	1.307	1.311
		Dolos $\downarrow$	0.274	0.256	0.265	0.258	0.251	0.240	0.234	0.216	0.216	0.215	0.199
CodeLlama (s=90)		Metrics	$\alpha = 1.0$	$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.7$	$\alpha = 0.6$	$\alpha = 0.5$	$\alpha = 0.4$	$\alpha = 0.3$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0$
Standard performance		pass@1 $\uparrow$	0.286	0.290	0.294	0.298	0.294	0.291	0.288	0.283	0.276	0.268	0.257
		pass@10 $\uparrow$	0.654	0.652	0.664	0.664	0.653	0.658	0.648	0.639	0.630	0.620	0.601
		pass@100 $\uparrow$	0.878	0.878	0.884	0.890	0.870	0.887	0.883	0.878	0.889	0.868	0.873
Copyright Infringement		EM $\downarrow$	23.430	21.692	19.383	15.928	11.726	7.433	4.540	3.051	2.270	1.715	1.311
		Dolos $\downarrow$	7.696	7.328	6.806	5.834	4.337	2.582	1.244	0.648	0.389	0.274	0.199

FIG. 9