We believe that more research is warranted to identify which of these approaches is most applicable in which cases, and how possible downsides of these alternatives can be mitigated in practice.

B   Appendix: Sample ambiguous inputs

Tables 4- 5 list inputs that are considered ambiguous by CICC in the B77 and HWU64 data sets respectively. Some inputs could refer to multiple intents whereas some other inputs could be considered out-of-scope.

#	Utterance	Label	Prediction Set
1	what is the matter?	direct debit payment not recognised	activate my card, age limit, balance not updated after bank transfer, balance not updated after cheque or cash deposit, beneficiary not allowed, cancel transfer, card arrival, card delivery estimate, card not working, card swallowed, cash withdrawal not recognised, change pin, compromised card, contactless not working, country support, declined card payment, declined transfer, direct debit payment not recognised, exchange rate, failed transfer, get physical card, lost or stolen card, lost or stolen phone, pending card payment, pending cash withdrawal, pending transfer, pin blocked, Refund not showing up, reverted card payment?, terminate account, top up failed, top up reverted, transaction charged twice, transfer not received by recipient, transfer timing, unable to verify identity, why verify identity, wrong amount of cash received,
2	Can I choose when my card is delivered?	card delivery estimate	activate my card, card about to expire, card acceptance, <u>card arrival</u>, <u>card delivery estimate</u>, change pin, contactless not working, country support, <u>get physical card</u>, getting spare card, getting virtual card, lost or stolen card, <u>order physical card</u>, supported cards and currencies, top up by bank transfer charge, top up by card charge, visa or mastercard
3	My contanctless has stopped working	contactless not working	activate my card, apple pay or google pay, automatic top up, beneficiary not allowed, cancel transfer, <u>card not working</u>, card payment wrong exchange rate, <u>contactless not working</u>, declined card payment, disposable card limits, failed transfer, get disposable virtual card, get physical card, pending top up, pin blocked, top up failed, top up reverted, topping up by card, virtual card not working, visa or mastercard, wrong exchange rate for cash withdrawal
4	I misplaced my card and I dont know where the last place is where I used the card last. Can you look at my account and tell me the last place I used the card?	lost or stolen card	activate my card, atm support, card acceptance, card linking, card swallowed, cash withdrawal not recognised, <u>compromised card</u>, <u>lost or stolen card</u>, lost or stolen phone, order physical card, <u>pin blocked</u>
5	Is my card denied anywhere?	card acceptance	<u>atm support</u>, <u>card acceptance</u>, card not working, card payment fee charged, <u>card swallowed</u>, compromised card, contactless not working, declined card payment, lost or stolen card, lost or stolen phone, order physical card, unable to verify identity, visa or mastercard

Table 4: A sample of prediction sets on B77 of size $> th$ of seven with marginal conformal prediction on BERT outputs. Plausible labels have been highlighted with <u>underscore</u>.

C   Appendix: LLM results

We here present a random sample of CQs on B77 and C150.

#	Utterance	Label	Prediction Set
1	olly	recommendation events	calendar set, general quirky, lists createoradd, music likeness, music query, play game, play music, play radio,
2	this song is too good	music likeness	audio volume mute, general affirm, general commandstop, general joke, general negate, lists remove, music dislikeness, <u>music likeness</u>
3	do i have to go to the gym	general quirky	calendar query, <u>general quirky</u>, lists query, <u>recommendation events</u>, recommendation locations, transport traffic, weather query
4	silently adjust	audio volume mute	<u>audio volume down</u>, <u>audio volume other</u>, <u>audio volume up</u>, iot hue lightchange, iot hue lightdim, iot hue lightup, <u>music settings</u>
5	how many times does it go	general quirky	datetime query, <u>general quirky</u>, <u>lists query</u>, qa factoid, qa maths, transport query, <u>transport traffic</u>
6	sports head lines please	news query	calendar set, general quirky, iot hue lightchange, music likeness, news query, qa factoid, social post, weather query
7	read that back	play audiobook	email addcontact, email query, email querycontact, email sendemail, general quirky, lists createoradd, music likeness, play audiobook, play music, social post,
8	i don't want to hear any more songs of that type	<u>music dislikeness</u>	audio volume mute, calendar remove, general commandstop, iot wemo off, lists remove, music dislikeness, music likeness
9	check celebrity wiki	general quirky	email query, <u>general quirky</u>, <u>lists query</u>, news query, <u>qa factoid</u>, social post, social query
10	Get all availables	lists query	email addcontact, email query, email querycontact, email sendemail, social post, social query, takeaway order,
11	rating	music likeness	cooking recipe, general quirky, lists createoradd, lists query, music likeness, music query, qa definition, qa factoid,
12	take me to mc donalds	transport query	play game, play podcasts, recommendation events, recommendation locations, recommendation movies, takeaway order, takeaway query
13	search	qa factoid	email querycontact, general quirky, lists createoradd, lists query, music query, qa definition, qa factoid,
14	unmute	audio volume up	audio volume down, <u>audio volume mute</u>, audio volume up, iot wemo off, music settings, play radio, transport query, <u>transport traffic</u>
15	please unmute yourself	audio volume mute	alarm remove, audio volume down, <u>audio volume mute</u>, <u>audio volume up</u>, iot cleaning, iot wemo on, music settings, play game
16	what's the best day next week to go out for pizza	datetime query	<u>calendar query</u>, cooking recipe, general quirky, qa factoid, <u>recommendation events</u>, recommendation locations, <u>takeaway query</u>
17	i need a manger	general quirky	calendar set, cooking recipe, general quirky, lists createoradd, music likeness, play game, qa definition, qa factoid, social post,
18	assistant shuffle entire library	play music	iot cleaning, iot hue lightchange, lists createoradd, <u>music settings</u>, play audiobook, play game, <u>play music</u>
19	put the disco lights on	iot hue lighton	alarm remove, iot cleaning, <u>iot hue lightchange</u>, iot hue lightoff, iot hue lighton, iot hue lightup, iot wemo on
20	hello how are you today	general greet	<u>general greet</u>, general praise, <u>general quirky</u>, play radio, recommendation events, recommendation locations, recommendation movies
21	where does tar work currently	email querycontact	cooking recipe, <u>email querycontact</u>, general quirky, lists query, qa definition, recommendation locations, takeaway query
22	can you pull up jeff	email querycontact	general quirky, iot cleaning, news query, play game, play music, play radio, social post
23	show me theatre where pirates of caribbean is playing	recommendation movies	general quirky, play game, play music, play podcasts, recommendation events, recommendation locations, <u>recommendation movies</u>
24	spot	play podcasts	calendar set, music likeness, music query, play music, play radio, qa factoid, weather query,
25	anything i need to know	general quirky	<u>calendar query</u>, calendar set, <u>email query</u>, email querycontact, <u>general quirky</u>, lists query, <u>news query</u>, <u>social query</u>
26	give me some good sounds	play music	audio volume down, audio volume up, music likeness, <u>music query</u>, music settings, <u>play music</u>, <u>play radio</u>
27	i don't like it	music dislikeness	calendar remove, general commandstop, <u>general dontcare</u>, general joke, <u>general negate</u>, lists remove, <u>music dislikeness</u>, music likeness, social post
28	what are you	general quirky	audio volume down, audio volume mute, audio volume up, music query, music settings, qa definition, recommendation locations, recommendation movies
29	i don't like this one	general quirky	<u>audio volume mute</u>, calendar remove, general joke, <u>general negate</u>, <u>iot wemo off</u>, <u>lists remove</u>, <u>music dislikeness</u>, social post

Table 5: Prediction sets on ACID of size $> th$ of seven with marginal conformal prediction on BERT outputs. Plausible labels have been highlighted with <u>underscore</u>.

Utterance	Prediction Set	LLM output
I hope you can help me. My account has been compromised from a different town that I haven't been to. Can you please place a hold on my account until I can make it to the bank?	cash withdrawal not recognised, compromised card	I understand you are having issues with your account. Was your account compromised or your card not recognised for cash withdrawal?
Is my card denied anywhere?	card not working, declined card payment	I understand you are asking about your card. Was it declined or not working?
Money that I have transferred hasn't arrived yet	balance not updated after bank transfer, transfer not received by recipient	I understand you are waiting for a transfer to be updated in your balance. Was the transfer not received by you or the recipient?
How long will the transfer take?	balance not updated after bank transfer, pending transfer, transfer timing	I understand you are asking about transfer timing. Was your transfer pending, not updated or not received?
What should I do with my stuck ATM?	card swallowed, declined cash withdrawal	I understand you have a stuck ATM. Was it your card that got stuck or was your cash withdrawal declined?

Table 6: Sample of LLM outputs on the B77 dataset where the user input has been made part of the prompt.