```
incr = bound / 10000
token = text.split()[-1]
title = f"GPT-J: p*,h({token}) - p*({token})"
color_schema = "PuOr"

utility.plot_results(prob, x, y, -bound, bound, incr, title, color_schema)

prob, access = ten_examples[i][1]
text, tkens, num_layers, _, _ = access
x = num_layers
y = tkens
bound = max(torch.max(prob).item(), abs(torch.min(prob).item()))
incr = bound / 10000
token = text.split()[-1]
title = f"GPT2-XL: p*,h({token}) - p*({token})"
color_schema = "PuOr"

utility.plot_results(prob, x, y, -bound, bound, incr, title, color_schema)
```


GPT-J: p*,h(actor) - p*(actor)

GPT2-XL: p*,h(actor) - p*(actor)

GPT-J: p*,h(Warsaw) - p*(Warsaw)

GPT2-XL: p*,h(Warsaw) - p*(Warsaw)

GPT-J: p*,h(London) - p*(London)

4

GPT2-XL: p*,h(London) - p*(London)

GPT-J: p*,h(Microsoft) - p*(Microsoft)

GPT2-XL: p*,h(Microsoft) - p*(Microsoft)

GPT-J: p*,h(BBC) - p*(BBC)

GPT2-XL: p*,h(BBC) - p*(BBC)

GPT-J: p*,h(Irish) - p*(Irish)

GPT2-XL: p*,h(Irish) - p*(Irish)

GPT-J: p*,h(Italian) - p*(Italian)

GPT2-XL: p*,h(Italian) - p*(Italian)

GPT-J: p*,h(Russian) - p*(Russian)

9

GPT2-XL: p*,h(Russian) - p*(Russian)

GPT-J: p*,h(Spanish) - p*(Spanish)

GPT2-XL: p*,h(Spanish) − p*(Spanish)

GPT-J: p*,h(Asia) − p*(Asia)

GPT2-XL: p*,h(Asia) - p*(Asia)