组会汇报

陈钶杰
专业:计算数学

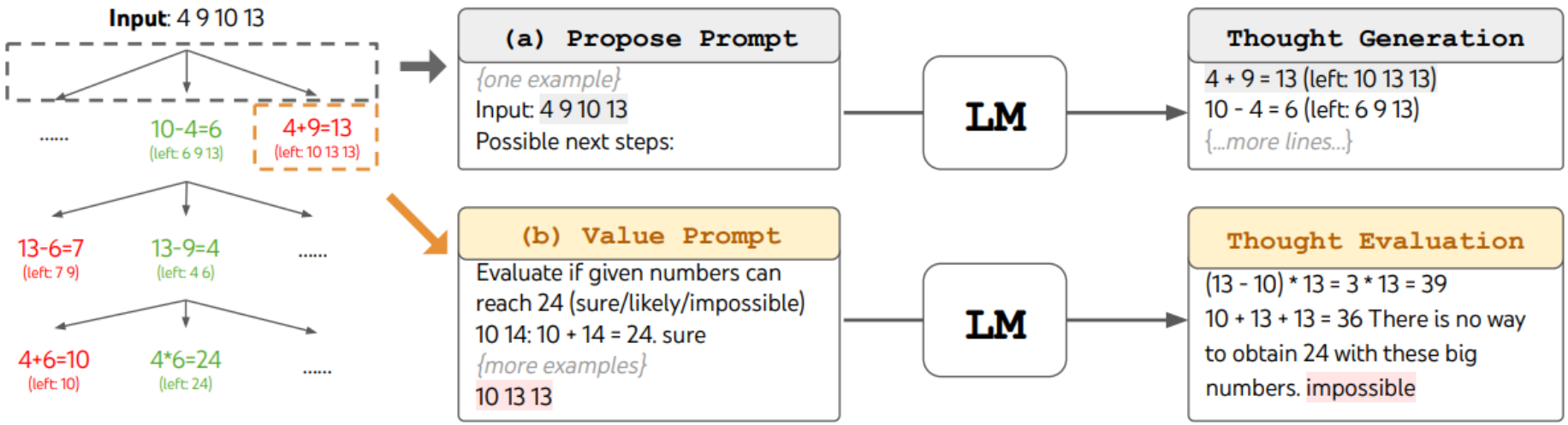
July 25, 2023

目录

1. Tree of Thought的实现思路:

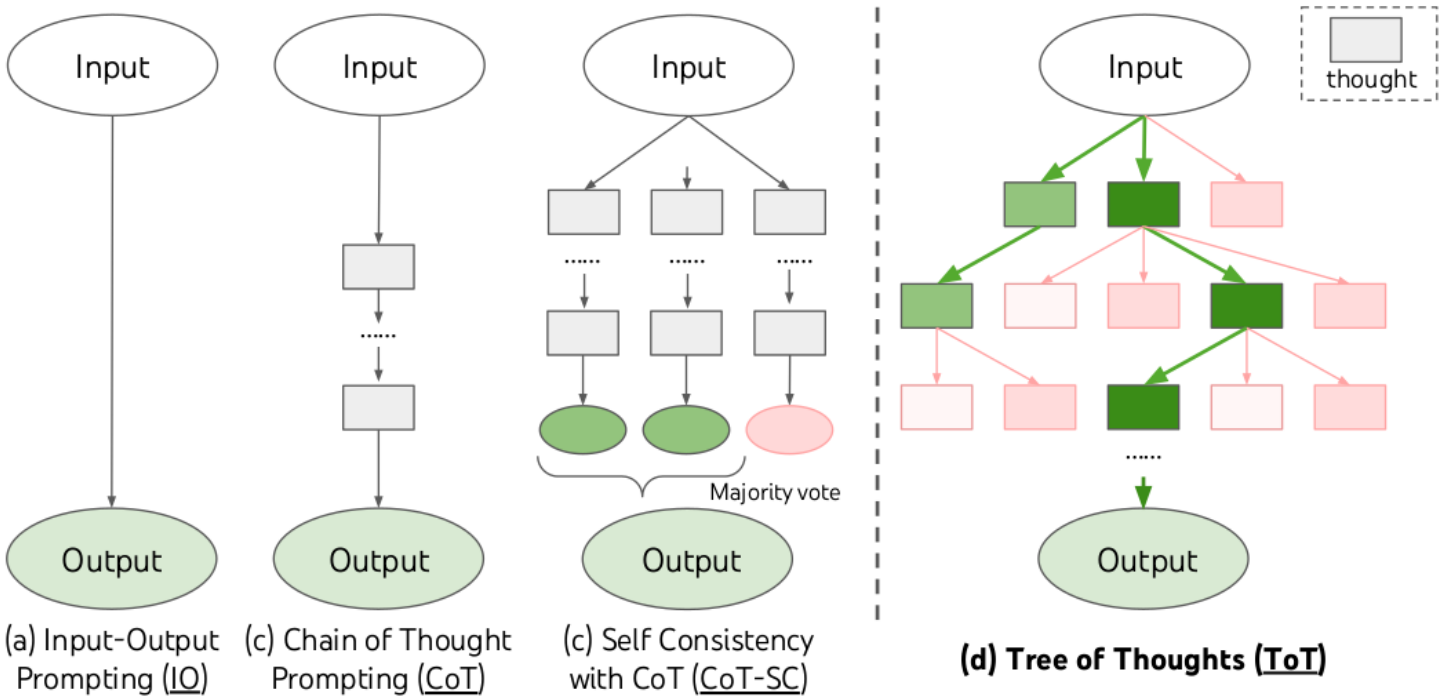
2. 代码调试
   - 使用特殊的提示的结果
   - 添加中间步骤的输出后的结果
   - 实验结果分析
   - 下一步的计划

Tree of Thoughts方法



- 用一个算24点的例子,来说明具体使用思维树.
  1. 思维分解:问题的提问方式可能一样(都是一句话),但是问题类型可以进行细分,比如算24点就是运算一行方程.
  2. 思维生成:生成多种可能,比如像上图中的多行thoughts.
  3. 思维评估:对生成的thoughts进行可能性分析.
  4. 搜索算法:选择合适的算法进行最优结果的搜索.比如这里选择BFS.

思维树的来源



- 简单输入输出.
- 除了输入输出以外,添加了中间步骤,以及在此基础上加了一些变式.
- 将中间步骤变成一种思维树的形式.

具体的中间步骤解释

- steps
  - step:0
    "x":"4 5 6 10"
    "ys":"之前有的组合"
    "new_ys":"新生成的组合"
    "values":"计算新生成的组合中评估值"
    "select_new_ys":"取values最高的5个作为下一步的ys"
  - :
  - step:3
    "x":"4 5 6 10"
    "ys":"可能的组合"
    "new_ys":"新生成的组合"
    "values":"计算新生成的组合中评估值"
    "select_new_ys":"取values最高的5个作为下一步的ys"

将现有的输出改成特殊提示的输出

- 原来的输入,输出
  1. input:预测长度为n的序列x_n后m结果?
  2. output:$x_{n+1}, \ldots, x_{n+m}$
- 修改后的输入,输出
  1. input:预测长度为n的序列x_n后m结果?
  2. output:

     step1:$x_{n-m}, \ldots, x_{n+1}$

     $\vdots$

     step(m-1):$x_n, \ldots, x_{n+m-1}$

     step(m):$x_{n+1}, \ldots, x_{n+m}$

     a

使用chatglm2其他提示的见过和没见过任务上的测试1


```
    {
        "content": "For a test Time-series, from the existing $16$ elements,predict the last column of future $8$ elements?
dim:4355,4356,4357,4358,4359,4360,4361,4362,4363,4364,4365,4366,4367,4368,4369,4370,dim:0.6835,0.9835,0.3793,-0.5737,-0.99
92,-0.5061,0.4523,0.9949,0.6227,-0.322,-0.9706,-0.7269,0.1851,0.927,0.8165,-0.04",
        "summary": "-0.8647,-0.8898,-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819"
    },
```

```
{"labels": "-0.8647,-0.8898,-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,", "predict": "- - - -=- -----"}
{"labels": "-0.8898,-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,", "predict": "- - -=-"}
{"labels": "-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,", "predict": "- - - -=-"}
{"labels": "0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984,-", "predict": "- -=-"}
{"labels": "0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984,-0.4602,-", "predict": "- - - -----"}
{"labels": "0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984,-0.4602,-0.9957,-", "predict": "- - - -=-"}
{"labels": "-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984,-0.4602,-0.9957,-0.6158,", "predict": "- - -=- -----"}
{"labels": "-0.9819,-0.3711,0.5809,0.9988,0.4984,-0.4602,-0.9957,-0.6158,0.330", "predict": "- - -=-"}
{"labels": "-0.3711,0.5809,0.9988,0.4984,-0.4602,-0.9957,-0.6158,0.3303,0.972", "predict": "- -=-=- - - - - -"}
{"labels": "0.5809,0.9988,0.4984,-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.720", "predict": "- - -=-"}
{"labels": "0.9988,0.4984,-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.193", "predict": "- - -=-"}
{"labels": "0.4984,-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.930", "predict": "- - -=-=-"}
{"labels": "-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.9302,-0.811", "predict": "- - -=- - - - - - - -000 - - - - -
- - - - -=- - -.. - - - - - - - - - - - - - - - - - - -)-)- - - - -"}
{"labels": "-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.9302,-0.8114,0.053", "predict": "- - -=-"}
{"labels": "-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.9302,-0.8114,0.0534,0.869", "predict": "- - -=-"}
{"labels": "0.3303,0.9727,0.7208,-0.1938,-0.9302,-0.8114,0.0534,0.8691,0.885", "predict": "- - - -=-"}
{"labels": "0.9727,0.7208,-0.1938,-0.9302,-0.8114,0.0534,0.8691,0.8858,0.088", "predict": "- -=-"}
{"labels": "0.7208,-0.1938,-0.9302,-0.8114,0.0534,0.8691,0.8858,0.088,-0.7907", "predict": "- - - - -----"}
{"labels": "-0.1938,-0.9302,-0.8114,0.0534,0.8691,0.8858,0.088,-0.7907,-0.9424", "predict": "- - - - - - - - - -"}
{"labels": "-0.9302,-0.8114,0.0534,0.8691,0.8858,0.088,-0.7907,-0.9424,-0.2277", "predict": "- - - - - - - - - - -000 - - - - -
- - - - -=- - -.. - - - - - - - - - - - - - - - - - - -"}
{"labels": "-0.8114,0.0534,0.8691,0.8858,0.088,-0.7907,-0.9424,-0.2277,0.6963", "predict": "- - - - -----"}
{"labels": "0.0534,0.8691,0.8858,0.088,-0.7907,-0.9424,-0.2277,0.6963,0.9802", "predict": "- -=-"}
{"labels": "0.8691,0.8858,0.088,-0.7907,-0.9424,-0.2277,0.6963,0.9802,0.3628", "predict": "- - - -=- - - - - - -000 - - - - -
```

```
{"labels": "0.4984,-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938", "predict": "- - - -=-"}
{"labels": "-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.9302", "predict": "- - - - -=-"}
{"labels": "-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.9302,-0.8114", "predict": "- - - -=- -----"}
{"labels": "-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.9302,-0.8114,0.0534", "predict": "- - -=-"}
{"labels": "0.3303,0.9727,0.7208,-0.1938,-0.9302,-0.8114,0.0534,0.8691", "predict": "- - - -=-"}
{"labels": "0.9727,0.7208,-0.1938,-0.9302,-0.8114,0.0534,0.8691,0.8858", "predict": "- -=-"}
{"labels": "0.7208,-0.1938,-0.9302,-0.8114,0.0534,0.8691,0.8858,0.088", "predict": "- - - -----"}
{"labels": "-0.1938,-0.9302,-0.8114,0.0534,0.8691,0.8858,0.088,-0.7907", "predict": "- - - -=-"}
{"labels": "-0.9302,-0.8114,0.0534,0.8691,0.8858,0.088,-0.7907,-0.9424", "predict": "- - -=- -----"}
{"labels": "-0.8114,0.0534,0.8691,0.8858,0.088,-0.7907,-0.9424,-0.2277", "predict": "- - -=-"}
{"labels": "0.0534,0.8691,0.8858,0.088,-0.7907,-0.9424,-0.2277,0.6963", "predict": "- -=-=- - - - - -"}
{"labels": "0.8691,0.8858,0.088,-0.7907,-0.9424,-0.2277,0.6963,0.9802", "predict": "- - -=-"}
{"labels": "0.8858,0.088,-0.7907,-0.9424,-0.2277,0.6963,0.9802,0.3628", "predict": "- - -=-"}
{"labels": "0.088,-0.7907,-0.9424,-0.2277,0.6963,0.9802,0.3628,-0.5881", "predict": "- - -=-=-"}
```

使用chatglm2其他提示的见过和没见过任务上的测试2


```
  {
    "content": "For a test Time-series, from the existing $8$ elements,predict the last column of future $4$ elements? 4365,-0.9706,#,4366,-0.7269,#,4367
,0.1851,#,4368,0.927,#,4369,0.8165,#,4370,-0.0446,#,4371,-0.8647,#,4372,-0.8898",
    "summary": "-0.0968,0.7852,0.9453,0.2363"
  },
```

```
{"labels": "-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809", "predict": "-    - - - - - -,- ----:."}
{"labels": "0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988", "predict": "- - - - - - -"}
{"labels": "0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984", "predict": "- - - - - -"}
{"labels": "0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984,-0.4602", "predict": "- - - - - - -"}
{"labels": "-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984,-0.4602,-0.9957", "predict": "-    - - - - -=- -----0"}
{"labels": "-0.9819,-0.3711,0.5809,0.9988,0.4984,-0.4602,-0.9957,-0.6158", "predict": "-   - - - - - - -,- --- ---"}
{"labels": "-0.3711,0.5809,0.9988,0.4984,-0.4602,-0.9957,-0.6158,0.3303", "predict": "-    - - - - - -"}
{"labels": "0.5809,0.9988,0.4984,-0.4602,-0.9957,-0.6158,0.3303,0.9727", "predict": "- - - - - -"}
{"labels": "0.9988,0.4984,-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.7208", "predict": "- - - - - -,- ---"}
{"labels": "0.4984,-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938", "predict": "- -    - -,- --"}
{"labels": "-0.4602,-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.9302", "predict": "-    - - - - - -=- ---0,-"}
{"labels": "-0.9957,-0.6158,0.3303,0.9727,0.7208,-0.1938,-0.9302,-0.8114", "predict": "-    - -    - -"}
```

```
{"labels": "-0.5061,0.4523,0.9949,0.6227,-0.322,-0.9706,-0.7269,0.1851,0.927,", "predict": "-    - - - - - -,- ----:."}
{"labels": "0.4523,0.9949,0.6227,-0.322,-0.9706,-0.7269,0.1851,0.927,0.8165,-", "predict": "-   - - - - - - -"}
{"labels": "0.9949,0.6227,-0.322,-0.9706,-0.7269,0.1851,0.927,0.8165,-0.0446,-", "predict": "-    - - - - - -"}
{"labels": "0.6227,-0.322,-0.9706,-0.7269,0.1851,0.927,0.8165,-0.0446,-0.8647,-", "predict": "- - - - - - -"}
{"labels": "-0.322,-0.9706,-0.7269,0.1851,0.927,0.8165,-0.0446,-0.8647,-0.8898,-", "predict": "- - - - -=- -----0"}
{"labels": "-0.9706,-0.7269,0.1851,0.927,0.8165,-0.0446,-0.8647,-0.8898,-0.0968", "predict": "- - - - - -,- --- ---"}
{"labels": "-0.7269,0.1851,0.927,0.8165,-0.0446,-0.8647,-0.8898,-0.0968,0.7852", "predict": "-    - - - - - -"}
{"labels": "0.1851,0.927,0.8165,-0.0446,-0.8647,-0.8898,-0.0968,0.7852,0.9453", "predict": "-    - - - - - -"}
{"labels": "0.927,0.8165,-0.0446,-0.8647,-0.8898,-0.0968,0.7852,0.9453,0.2363", "predict": "-   - - - - - -,- ---"}
{"labels": "0.8165,-0.0446,-0.8647,-0.8898,-0.0968,0.7852,0.9453,0.2363,-0.69,-", "predict": "-    - -    - -,- --"}
{"labels": "-0.0446,-0.8647,-0.8898,-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819,-", "predict": "- - - - - -=- ---0,-"}
{"labels": "-0.8647,-0.8898,-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,", "predict": "-    - -    - -"}
{"labels": "-0.8898,-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,", "predict": "- - - - - - -"}
{"labels": "-0.0968,0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,", "predict": "-   - - -  ,- -"}
{"labels": "0.7852,0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984,-", "predict": "-   - - -    - ----------------"}
{"labels": "0.9453,0.2363,-0.69,-0.9819,-0.3711,0.5809,0.9988,0.4984,-0.4602,-", "predict": "-    - - - - - ----------------"}
```

添加部分步骤的提示结果


```
{
    "content": "For a test Time-series, from the existing $32$ elements,predict the last column of future $16$ elements\
? dim:4333,4334,4335,4336,4337,4338,4339,4340,4341,4342,4343,4344,4345,4346,4347,4348,4349,4350,4351,4352,4353,4354,4355,43\
56,4357,4358,4359,4360,4361,4362,4363,4364,dim:-0.7359,0.1721,0.9219,0.8241,-0.0313,-0.858,-0.8958,-0.11,0.7769,0.9496,0.24\
92,-0.6803,-0.9843,-0.3834,0.57,0.9994,0.5099,-0.4484,-0.9944,-0.6262,0.3178,0.9695,0.7299,-0.1808,-0.9253,-0.8191,0.0402,0\
.8625,0.8918,0.1012,-0.7825,-0.9468,dim:-0.9665,0.6107,-0.9081,-0.8043,0.955,-0.7237,-0.7759,0.8589,-0.9709,-0.9928,-0.0532\
,-0.4287,0.2008,-0.9138,0.979,0.4234,0.9981,-0.9913,-0.225,-0.8497,0.6148,-0.7398,-0.3218,0.8215,0.389,0.5823,0.2937,0.7049\
,0.6602,0.4633,0.3103,-0.0289,dim:-0.677,-0.9851,-0.3874,0.5664,0.9995,0.5137,-0.4444,-0.9939,-0.6296,0.3136,0.9685,0.7329,\
-0.1764,-0.9236,-0.8216,0.0358,0.8603,0.8938,0.1056,-0.7797,-0.9482,-0.2449,0.6835,0.9835,0.3793,-0.5737,-0.9992,-0.5061,0.\
4523,0.9949,0.6227,-0.3",
    "summary": "step1:0.8938, 0.1056, -0.7797, -0.9482, -0.2449, 0.6835, 0.9835, 0.3793, -0.5737, -0.9992, -0.5061, 0.4\
523, 0.9949, 0.6227, -0.322,-0.9706,\n,step2:0.1056, -0.7797, -0.9482, -0.2449, 0.6835, 0.9835, 0.3793, -0.5737, -0.9992, -\
0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0.9706, -0.7269,\n,step3:-0.7797, -0.9482, -0.2449, 0.6835, 0.9835, 0.3793, -0.5737\
, -0.9992, -0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0.9706, -0.7269, 0.1851,\n,step4:-0.9482, -0.2449, 0.6835, 0.9835, 0.37\
93, -0.5737, -0.9992, -0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0.9706, -0.7269, 0.1851, 0.927,\n,step5:-0.2449, 0.6835, 0.9\
835, 0.3793, -0.5737, -0.9992, -0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0.9706, -0.7269, 0.1851, 0.927, 0.8165,\n,step6:0.6\
835, 0.9835, 0.3793, -0.5737, -0.9992, -0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0.9706, -0.7269, 0.1851, 0.927, 0.8165, -0.\
0446,\n,step7:0.9835, 0.3793, -0.5737, -0.9992, -0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0.9706, -0.7269, 0.1851, 0.927, 0.\
8165, -0.0446, -0.8647,\n,step8:0.3793, -0.5737, -0.9992, -0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0.9706, -0.7269, 0.1851,\
 0.927, 0.8165, -0.0446, -0.8647, -0.8898,\n,step9:-0.5737, -0.9992, -0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0.9706, -0.72\
69, 0.1851, 0.927, 0.8165, -0.0446, -0.8647, -0.8898, -0.0968,\n,step10:-0.9992, -0.5061, 0.4523, 0.9949, 0.6227, -0.322,-0\
.9706, -0.7269, 0.1851, 0.927, 0.8165, -0.0446, -0.8647, -0.8898, -0.0968, 0.7852,\n,step11:-0.5061, 0.4523, 0.9949, 0.6227\
, -0.322,-0.9706, -0.7269, 0.1851, 0.927, 0.8165, -0.0446, -0.8647, -0.8898, -0.0968, 0.7852, 0.9453,\n,step12:0.4523, 0.99\
49, 0.6227, -0.322,-0.9706, -0.7269, 0.1851, 0.927, 0.8165, -0.0446, -0.8647, -0.8898, -0.0968, 0.7852, 0.9453, 0.2363,\n,s\
tep13:0.9949, 0.6227, -0.322,-0.9706, -0.7269, 0.1851, 0.927, 0.8165, -0.0446, -0.8647, -0.8898, -0.0968, 0.7852, 0.9453, 0\
.2363, -0.69,\n,step14:0.6227, -0.322,-0.9706, -0.7269, 0.1851, 0.927, 0.8165, -0.0446, -0.8647, -0.8898, -0.0968, 0.7852, \
0.9453, 0.2363, -0.69, -0.9819,\n,step15:-0.322,-0.9706, -0.7269, 0.1851, 0.927, 0.8165, -0.0446, -0.8647, -0.8898, -0.0968\
, 0.7852, 0.9453, 0.2363, -0.69, -0.9819, -0.3711,\n,result:-0.9706, -0.7269, 0.1851, 0.927, 0.8165, -0.0446, -0.8647, -0.8\
898, -0.0968, 0.7852, 0.9453, 0.2363, -0.69, -0.9819, -0.3711, 0.5809"
},
```



```
1{"labels": "step1:-0.7269, 0.1851, 0.927, 0.8165, -0.0446,-0.8647,\n,step2:0.1851, 0", "predict": "step0000000006600000.0..0.000000000.0000000000..00.0.0000.....0.....}
2{"labels": "step1:0.1851, 0.927, 0.8165, -0.0446, -0.8647,-0.8898,\n,step2:0.927, 0.", "predict": "step0800000006600000..0000000000..00000000000..000....0....00.......}
3{"labels": "step1:0.927, 0.8165, -0.0446, -0.8647, -0.8898,-0.0968,\n,step2:0.8165, -0", "predict": "step00000000060.0000...0..00000...0.000.0000..000.0000000...00......}
4{"labels": "step1:0.8165, -0.0446, -0.8647, -0.8898, -0.0968,0.7852,\n,step2:-0.0446, -", "predict": "step00000000066..000....0.00000.000000000.00..00...0..000.0..0.......}
5{"labels": "step1:-0.0446, -0.8647, -0.8898, -0.0968, 0.7852,0.9453,\n,step2:-0.8647, -", "predict": "step0800000066600000...000000000.00..0..0.0000.00.000000....00......}
6{"labels": "step1:-0.8647, -0.8898, -0.0968, 0.7852, 0.9453,0.2363,\n,step2:-0.8898, -", "predict": "step0000000006600000..0...000000.00000.00000.0...0.00.000..000......}
7{"labels": "step1:-0.8898, -0.0968, 0.7852, 0.9453, 0.2363,-0.69,\n,step2:-0.0968, 0.", "predict": "step000000000600..0.0..0..00000000.0.00.0.0000..0..0.0..00.........}
8{"labels": "step1:-0.0968, 0.7852, 0.9453, 0.2363, -0.69,-0.9819,\n,step2:0.7852, 0.", "predict": "step000000000600.00..00.00000000..00000.0000.0..0.00..00...0.0......}
```

实验结论

- chatglm2的训练结果不如chatglm.之后测试可能还需要用chatglm来做
- 可能是因为

下一步计划

- 部署llama2,使其能够像chatglm一样进行类似训练和评估.
- 通过其他方式在chatglm上添加中间步骤看看效果如何?
- 提高模型的泛化能力
  1. 添加思维树的方式来提高泛化能力

谢谢老师和同学们的聆听!