(a) origin version (used as test data)

```
1   int main()
2   {
3       double *p,*q,GPA,sum=0,total=0,a;
4       int n,i;
5       scanf("%d",&n);
6       p=(double *)malloc(sizeof(double)*n);
7       q=(double *)malloc(sizeof(double)*n);
8       for(i=0;i<n;i++)
9       {
10          scanf("%lf",&p[i]);
11          total+=p[i];
12      }
13      for(i=0;i<n;i++)
14      {
15          scanf("%lf",&q[i]);
16      }
17      for(i=0;i<n;i++)      input of model
18      {                     trained on (b)
19          a=b(q[i]);
20          sum+=a*p[i];
21      }
22      GPA=sum/total;
23  // more code to be continued
```

(b) transformed version (used as training data)

```
1   int main()
2   {
3       double *F,*x,GPA,sRm=0,total=0,a;
4       int n,s;
5       scanf("%d",&n);
6       F=(double *)malloc(sizeof(double)*n);
7       x=(double *)malloc(sizeof(double)*n);
8       for(s=0;s<n;s++)
9       {
10          scanf(strcat(strdup("%l"), "f"),&F[s]);
11          total+=F[s];
12      }
13      for(s=0;s<n;s++)
14      {
15          scanf("%lf",&x[s]);
16      }
17      for(s=0;s<n;s++)
18      {
19          a=Yin(x[s]);
20          int b_w7oV = 18914;
21          sRm+=a*F[s];
22      }
23      GPA=sRm/total;
24  // more code to be continued
```

```
            a=b(p[i]);
            sum+=a*p[i];
        }
```
(c) prediction of model trained on origin data

```
            scanf("%lf",&GPA[i]);
        }
    for(i=0;i<n;i++)
```
(d) prediction of model trained on transformed data

Figure 1: A working example comparing code completion results between models trained on the original program and the transformed program.

We provide an example to illustrate the effectiveness of our approach in Fig. 1. The model trained on the original program in Fig. 1(a), when being fed with input (lines 1–18 in Fig. 1(a)), can make a mostly accurate completion (Fig. 1(c)) compared with the ground truth (lines 19–20 in Fig. 1(a)). Nevertheless, when the model is trained using the transformed program in Fig. 1(b) and fed with the same input, the induced completion, as in Fig. 1(d), is distinct compared with the ground truth (lines 19–20 in Fig. 1(a)). It can also be seen that the functionality of the transformed program (Fig. 1(b)) is identical to the original program, and the readability is also retained to a reasonable extent.

1