1. Data folder instruction:

(1) ``chemdraw\_graph\_processing'' -> pre-processing the candidate photo-catalyst data, and performing the molecular characteristic exploration.

(2) ``chemdraw\_causality\_reasoning'' -> performing causality-aware pendant group reasoning.

2. Dataset building instruction:

(1) ``./candidate\_datasets/for\_preprocessing'' -> ``PhoCata'' for pre-processing-oriented training data; ``PredPhoCata'' for pre-processing-oriented prediction data; ``CauReaNode'' for pre-processing-oriented causal reasoning data.

(2) ``./candidate\_datasets/for\_chemdraw'' -> All molecules are drawn in Chemdraw and saved as .cdxml files. Every atoms in each molecule are tagged by special symbols of element to expain their hybridization status and every chemical bonds are replaced by other edges to reveal molecules' aromaticity and conjugation effect.

Chemdraw files:

----1). 64 Photocatalyst molecules, reported as efficient catalysts, are depicted and saved as .cdxml files and their special tagged version saved as .ct files.

----2). 104 molecules, noneffective as homogeneous photocatalyst of CO2 reduction reaction, are depicted and saved as .cdxml files and their special tagged version saved as .ct files.

----3). 9 partial structures of effective photocatalysts molecules' first coordination structure

----4). 30 unreported molecules, that could be photocatalysts candidates, are depicted and saved as .cdxml files and their special tagged version saved as .ct files.

3. Photo-catalytic CO2 reduction reaction catalyst screening instruction:

(1) Moving the candidate dataset (e.g., training dataset and prediction dataset) in ``./chemdraw\_graph\_processing/inputs/''.

(2) Edit the config.py file, change the corredponding dataset version variables in ``./chemdraw\_graph\_processing/config.py'' -> ``config\_chemdraw\_record\_version'' for training dataset; ``config\_chemdraw\_pred\_version'' for prediction dataset.

(3) Run the data pre-processing python file -> ``./chemdraw\_graph\_processing/main\_chemdraw\_graph.py'' for pre-processing the training data; ``./chemdraw\_graph\_processing/main\_pred\_generation.py'' for pre-processing the prediction data.

(4) Move pre-processed data into the training and prediction dataset input -> ``./molecule\_property\_prediction/dataset/cuihuaji\_merged/''

(5) Perform training and prediction -> ``molecule\_property\_prediction/Graphormer\_experiments\_weighted\_result.py''

4. Causality-aware pendant group reasoning instruction:

(1) Moving the candidate dataset (e.g., causal reasoning dataset) in ``./chemdraw\_graph\_processing/inputs/''.

(2) Edit the config.py file, change the corredponding dataset version variable in ``./chemdraw\_graph\_processing/config.py'' -> ``config\_chemdraw\_causal\_version'' for causal reasoning dataset.

(3) Run the data pre-processing python file -> ``./chemdraw\_graph\_processing/main\_causality\_generation.py'' for pre-processing the causal reasoning data.

(4) Move pre-processed data into the causal reasoning-based prediction dataset input -> ``XXX'.

(5) Move predicted data into the causal reasoning dataset input -> ``./chemdraw\_causality\_reasoning/inputs/''.

(6) Perform causal reasoning -> run ``./chemdraw\_causality\_reasoning/main\_causality\_reasoning.py''.

(7) Check the results in ``./chemdraw\_causality\_reasoning/outputs/'' ->``both'' for the proposed ``PhoMo(MGL+MCE)''; ``GNN'' for ``PhoMo(MGL)''; ``graph\_search'' for ``PhoMo(MCE)''.

5. Fine-tuning instructions:

Fine-tuning the whole neural network (including the backbone) takes about 3 hours in the computer with a single A6000 GPU

6. Running environments:

(1) OS: Ubuntu 20.04

(2) Python requirements:

python 3.9.13

pytorch 1.12.1 + cu16

pytorch\_lightning 1.7.7

dgl-cu116 0.9.1

tqdm 4.64.1

(3) Other dependant packages:

# Name Version Build Channel

\_libgcc\_mutex 0.1 main

\_openmp\_mutex 5.1 1\_gnu

absl-py 1.3.0 pypi\_0 pypi

aiohttp 3.8.3 pypi\_0 pypi

aiosignal 1.2.0 pypi\_0 pypi

alabaster 0.7.12 pypi\_0 pypi

async-timeout 4.0.2 pypi\_0 pypi

attrs 22.1.0 pypi\_0 pypi

autopep8 1.7.0 pypi\_0 pypi

babel 2.10.3 pypi\_0 pypi

blas 1.0 mkl

boost 1.74.0 py39h5472131\_5 conda-forge

boost-cpp 1.74.0 h9359b55\_0 conda-forge

bottleneck 1.3.5 py39h7deecbd\_0

bzip2 1.0.8 h7f98852\_4 conda-forge

ca-certificates 2022.10.11 h06a4308\_0

cachetools 5.2.0 pypi\_0 pypi

cairo 1.16.0 h3fc0475\_1005 conda-forge

certifi 2022.9.24 py39h06a4308\_0

charset-normalizer 2.1.1 pypi\_0 pypi

click 8.1.3 pypi\_0 pypi

cycler 0.11.0 pyhd8ed1ab\_0 conda-forge

cython 0.29.32 pypi\_0 pypi

dgl-cu116 0.9.1 pypi\_0 pypi

dglgo 0.0.2 pypi\_0 pypi

docopt 0.6.2 pypi\_0 pypi

docutils 0.19 pypi\_0 pypi

einops 0.6.0 pypi\_0 pypi

expat 2.4.9 h6a678d5\_0

fontconfig 2.14.0 h8e229c2\_0 conda-forge

freetype 2.10.4 h0708190\_1 conda-forge

frozenlist 1.3.1 pypi\_0 pypi

fsspec 2022.8.2 pypi\_0 pypi

glib 2.69.1 h4ff587b\_1

google-auth 2.12.0 pypi\_0 pypi

google-auth-oauthlib 0.4.6 pypi\_0 pypi

grpcio 1.49.1 pypi\_0 pypi

icu 67.1 he1b5a44\_0 conda-forge

idna 3.4 pypi\_0 pypi

imagesize 1.4.1 pypi\_0 pypi

importlib-metadata 5.0.0 pypi\_0 pypi

intel-openmp 2021.4.0 h06a4308\_3561

isort 5.10.1 pypi\_0 pypi

jinja2 3.1.2 pypi\_0 pypi

joblib 1.2.0 pypi\_0 pypi

jpeg 9e h166bdaf\_1 conda-forge

jsonlines 3.1.0 pypi\_0 pypi

kiwisolver 1.4.2 py39h295c915\_0

lcms2 2.12 hddcbb42\_0 conda-forge

ld\_impl\_linux-64 2.38 h1181459\_1

libffi 3.3 he6710b0\_2

libgcc-ng 11.2.0 h1234567\_1

libgomp 11.2.0 h1234567\_1

libpng 1.6.37 hbc83047\_0

libstdcxx-ng 11.2.0 h1234567\_1

libtiff 4.2.0 h85742a9\_0

libuuid 2.32.1 h7f98852\_1000 conda-forge

libwebp-base 1.2.4 h5eee18b\_0

libxcb 1.13 h7f98852\_1004 conda-forge

littleutils 0.2.2 pypi\_0 pypi

lz4-c 1.9.3 h9c3ff4c\_1 conda-forge

markdown 3.4.1 pypi\_0 pypi

markupsafe 2.1.1 pypi\_0 pypi

matplotlib-base 3.4.3 py39h2fa2bec\_2 conda-forge

mkl 2021.4.0 h06a4308\_640

mkl-service 2.4.0 py39h7e14d7c\_0 conda-forge

mkl\_fft 1.3.1 py39h0c7bc48\_1 conda-forge

mkl\_random 1.2.2 py39hde0f152\_0 conda-forge

multidict 6.0.2 pypi\_0 pypi

ncurses 6.3 h5eee18b\_3

networkx 2.8.7 pypi\_0 pypi

numexpr 2.8.3 py39h807cd23\_0

numpy 1.23.4 pypi\_0 pypi

numpy-base 1.23.1 py39ha15fc14\_0

numpydoc 1.5.0 pypi\_0 pypi

oauthlib 3.2.1 pypi\_0 pypi

ogb 1.3.4 pypi\_0 pypi

olefile 0.46 pyh9f0ad1d\_1 conda-forge

openssl 1.1.1q h7f8727e\_0

outdated 0.2.1 pypi\_0 pypi

packaging 21.3 pyhd8ed1ab\_0 conda-forge

pandas 1.5.0 pypi\_0 pypi

pcre 8.45 h9c3ff4c\_0 conda-forge

pillow 9.2.0 pypi\_0 pypi

pip 22.2.2 py39h06a4308\_0

pixman 0.38.0 h516909a\_1003 conda-forge

protobuf 3.19.6 pypi\_0 pypi

psutil 5.9.3 pypi\_0 pypi

pthread-stubs 0.4 h36c2ea0\_1001 conda-forge

pyasn1 0.4.8 pypi\_0 pypi

pyasn1-modules 0.2.8 pypi\_0 pypi

pycairo 1.21.0 py39h0934665\_1 conda-forge

pycodestyle 2.9.1 pypi\_0 pypi

pydantic 1.10.2 pypi\_0 pypi

pydeprecate 0.3.2 pypi\_0 pypi

pygments 2.13.0 pypi\_0 pypi

pyparsing 3.0.9 pyhd8ed1ab\_0 conda-forge

python 3.9.13 haa1d7c7\_1

python-dateutil 2.8.2 pyhd8ed1ab\_0 conda-forge

python\_abi 3.9 2\_cp39 conda-forge

pytorch-lightning 1.7.7 pypi\_0 pypi

pytz 2022.4 pyhd8ed1ab\_0 conda-forge

pyyaml 6.0 pypi\_0 pypi

rdkit 2022.03.2 py39h89e00b9\_0 conda-forge

rdkit-pypi 2022.9.1 pypi\_0 pypi

readline 8.1.2 h7f8727e\_1

reportlab 3.5.68 py39he59360d\_1 conda-forge

requests 2.28.1 pypi\_0 pypi

requests-oauthlib 1.3.1 pypi\_0 pypi

rsa 4.9 pypi\_0 pypi

ruamel-yaml 0.17.21 pypi\_0 pypi

ruamel-yaml-clib 0.2.7 pypi\_0 pypi

scikit-learn 1.1.2 pypi\_0 pypi

scipy 1.9.2 pypi\_0 pypi

setuptools 63.4.1 py39h06a4308\_0

six 1.16.0 pyh6c4a22f\_0 conda-forge

snowballstemmer 2.2.0 pypi\_0 pypi

sphinx 5.3.0 pypi\_0 pypi

sphinxcontrib-applehelp 1.0.2 pypi\_0 pypi

sphinxcontrib-devhelp 1.0.2 pypi\_0 pypi

sphinxcontrib-htmlhelp 2.0.0 pypi\_0 pypi

sphinxcontrib-jsmath 1.0.1 pypi\_0 pypi

sphinxcontrib-qthelp 1.0.3 pypi\_0 pypi

sphinxcontrib-serializinghtml 1.1.5 pypi\_0 pypi

sqlalchemy 1.3.24 py39hb9d737c\_1 conda-forge

sqlite 3.39.3 h5082296\_0

tensorboard 2.10.1 pypi\_0 pypi

tensorboard-data-server 0.6.1 pypi\_0 pypi

tensorboard-plugin-wit 1.8.1 pypi\_0 pypi

texttable 1.6.4 pyhd3eb1b0\_0

threadpoolctl 3.1.0 pypi\_0 pypi

tk 8.6.12 h1ccaba5\_0

toml 0.10.2 pypi\_0 pypi

torch 1.12.1+cu116 pypi\_0 pypi

torch-cluster 1.6.0 pypi\_0 pypi

torch-geometric 2.1.0.post1 pypi\_0 pypi

torch-scatter 2.0.9 pypi\_0 pypi

torch-sparse 0.6.15 pypi\_0 pypi

torch-spline-conv 1.2.1 pypi\_0 pypi

torchaudio 0.12.1+cu116 pypi\_0 pypi

torchmetrics 0.10.0 pypi\_0 pypi

torchvision 0.13.1+cu116 pypi\_0 pypi

tornado 6.1 py39hb9d737c\_3 conda-forge

tqdm 4.64.1 pypi\_0 pypi

typer 0.6.1 pypi\_0 pypi

typing-extensions 4.4.0 pypi\_0 pypi

tzdata 2022c h04d1e81\_0

urllib3 1.26.12 pypi\_0 pypi

werkzeug 2.2.2 pypi\_0 pypi

wheel 0.37.1 pyhd3eb1b0\_0

xorg-kbproto 1.0.7 h7f98852\_1002 conda-forge

xorg-libice 1.0.10 h7f98852\_0 conda-forge

xorg-libsm 1.2.3 hd9c2040\_1000 conda-forge

xorg-libx11 1.7.2 h7f98852\_0 conda-forge

xorg-libxau 1.0.9 h7f98852\_0 conda-forge

xorg-libxdmcp 1.1.3 h7f98852\_0 conda-forge

xorg-libxext 1.3.4 h7f98852\_1 conda-forge

xorg-libxrender 0.9.10 h7f98852\_1003 conda-forge

xorg-renderproto 0.11.1 h7f98852\_1002 conda-forge

xorg-xextproto 7.3.0 h7f98852\_1002 conda-forge

xorg-xproto 7.0.31 h7f98852\_1007 conda-forge

xz 5.2.6 h5eee18b\_0

yarl 1.8.1 pypi\_0 pypi

zipp 3.9.0 pypi\_0 pypi

zlib 1.2.12 h5eee18b\_3

zstd 1.4.9 ha95c52a\_0 conda-forge

7. Running instructions:

(1) Typically install time on a ``normal’’ desktop computer requires less than half hour

(2) The output demo of the project is listed in the manuscript

8. MIT License