Towards functorial chemistry

Completeness and universality for reaction representation

Leo Lobski

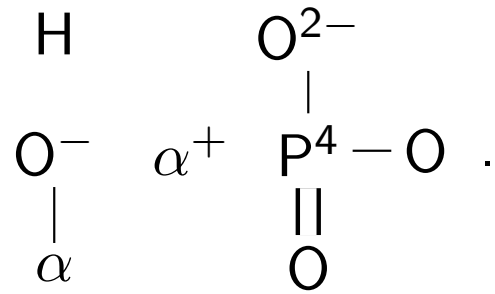
UCL

Birmingham Theory Seminar
15 March 2024

Outline

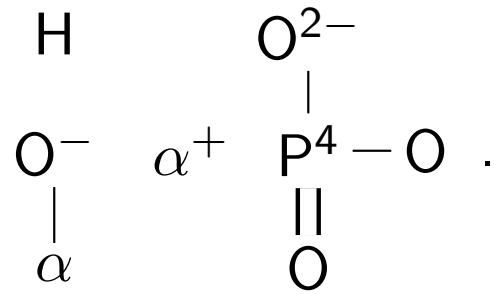
Chemical graphs and morphisms

Molecular entities are represented by labelled graphs:

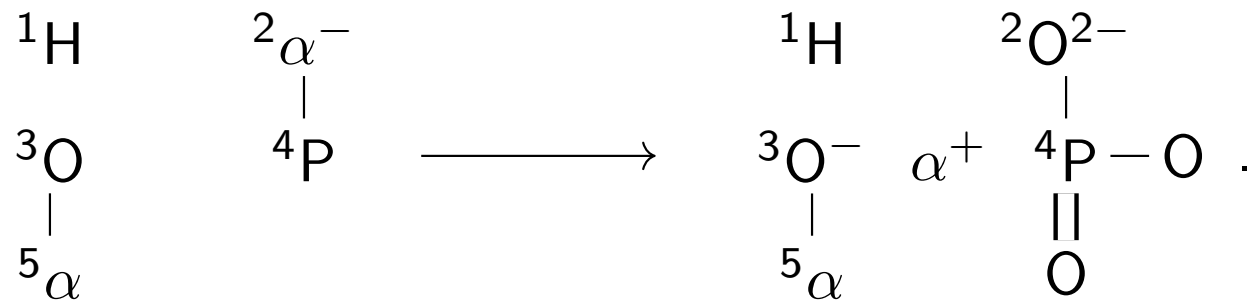


Chemical graphs and morphisms

Molecular entities are represented by labelled graphs:

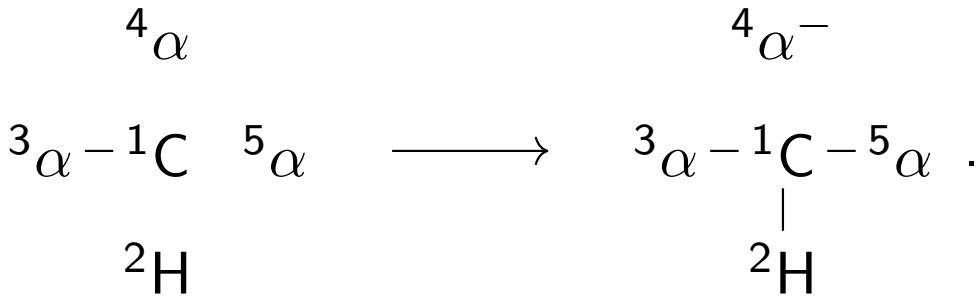


A morphism is a “matter non-decreasing” function:



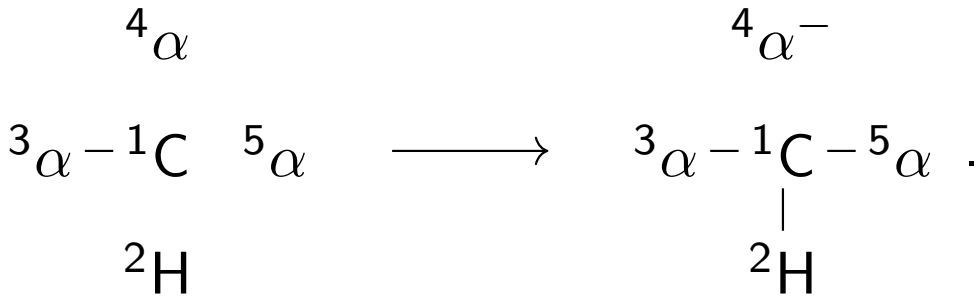
Vertex embeddings and matchings

- A *vertex embedding* is a bijective morphism which preserves the atom labels:

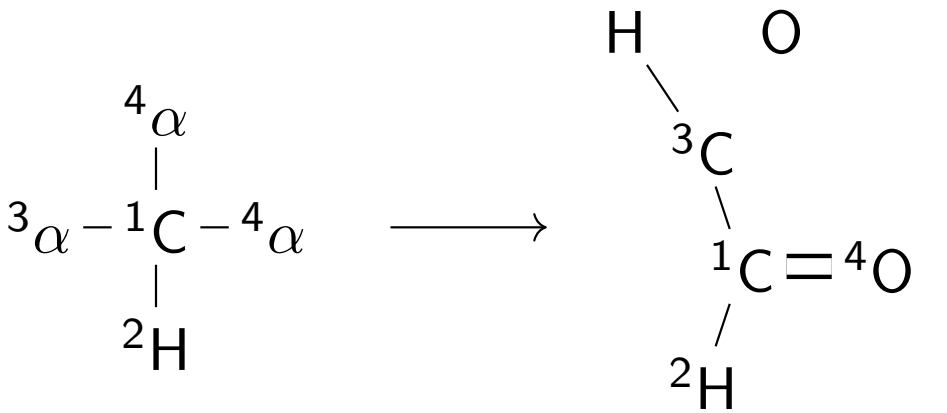


Vertex embeddings and matchings

- A *vertex embedding* is a bijective morphism which preserves the atom labels:

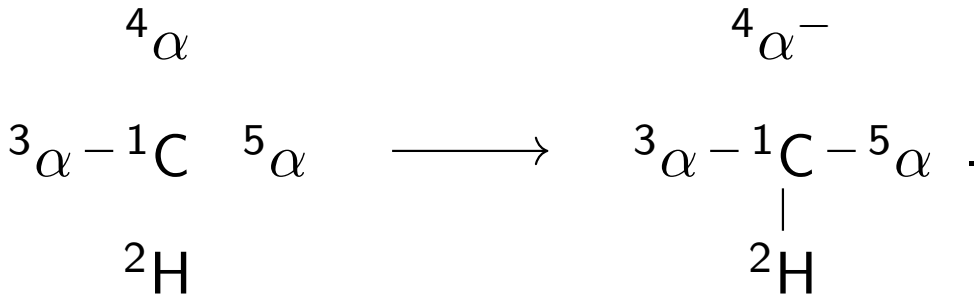


- A *matching* is a morphism which strictly preserves the charge and the number of bonds:

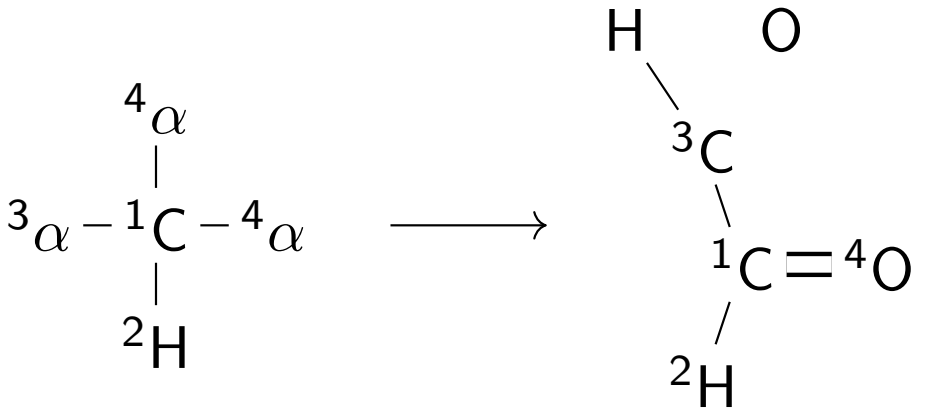


Vertex embeddings and matchings

- A *vertex embedding* is a bijective morphism which preserves the atom labels:



- A *matching* is a morphism which strictly preserves the charge and the number of bonds:

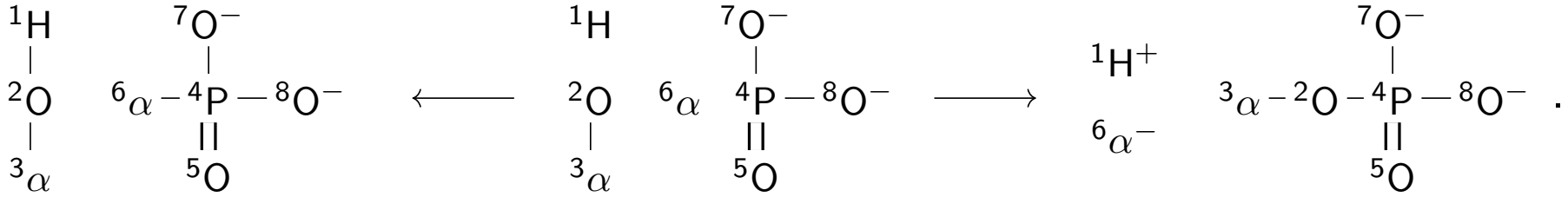


Theorem

The category of chemical graphs is $(\mathcal{E}, \mathcal{M})$*-adhesive.*

Reaction schemes

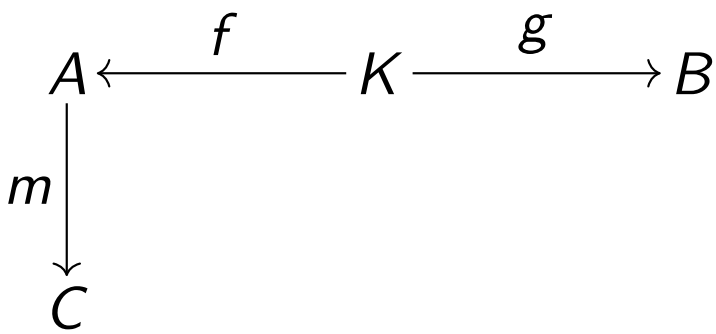
A *reaction scheme* is a terminal span of vertex embeddings whose boundary graphs have the same net charge and are valence complete:

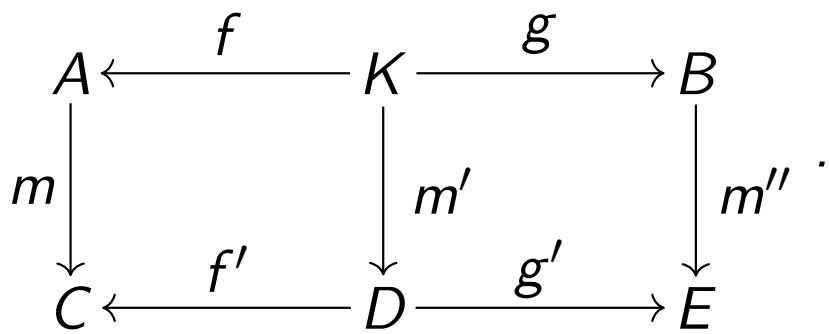


Reaction schemes

Proposition

The diagram below left, where the top span is a reaction scheme and $m \in \mathcal{M}$, can be uniquely completed to the one on the right, with $f', g' \in \mathcal{E}$ and $m', m'' \in \mathcal{M}$

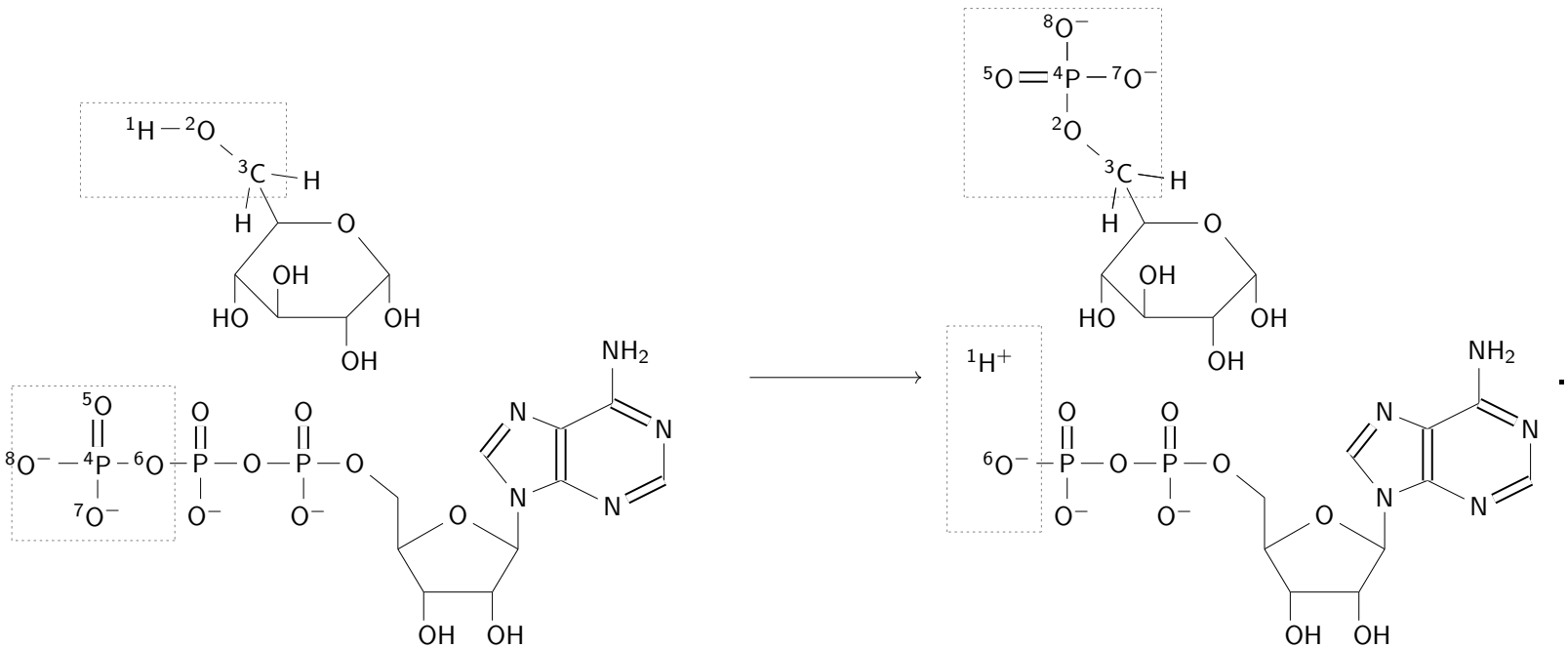




Moreover, if C is valence complete, then so is E.

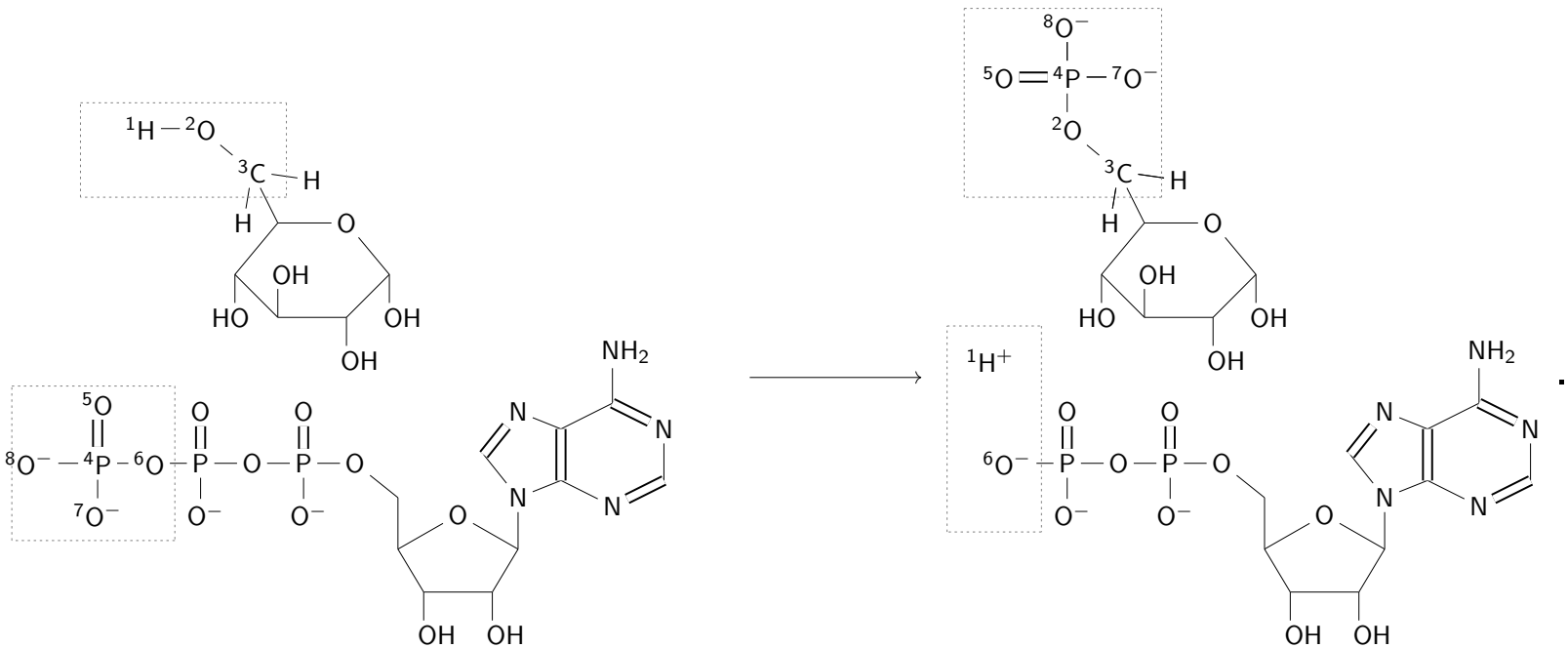
Reactions

A *reaction* between valence complete chemical graphs is a partial bijection such that the domain and the image have the same net charge, atom labels are preserved and boundary vertices are mapped to boundary vertices:



Reactions

A *reaction* between valence complete chemical graphs is a partial bijection such that the domain and the image have the same net charge, atom labels are preserved and boundary vertices are mapped to boundary vertices:

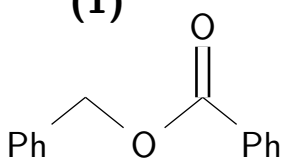


We denote the category of reactions by **React**.

Retrosynthetic analysis

(1) Start with the target molecule(s)

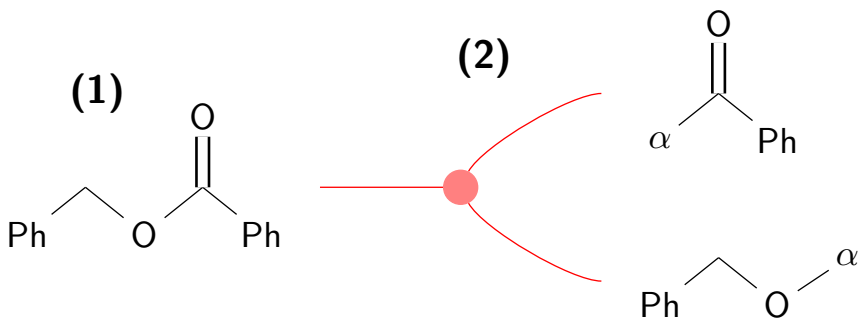
(1)



Retrosynthetic analysis

(1) Start with the target molecule(s)

(2) Cut the target along some bond, creating *synthons*

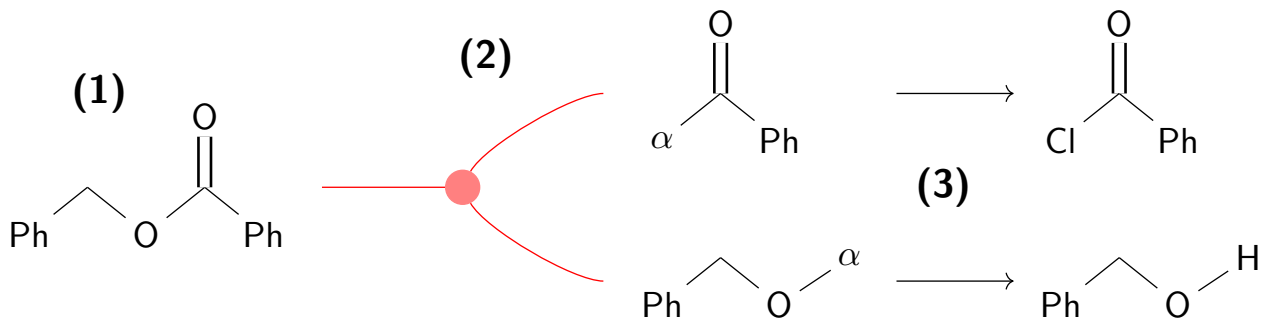


Retrosynthetic analysis

(1) Start with the target molecule(s)

(2) Cut the target along some bond, creating *synthons*

(3) Search for *synthetic equivalents*



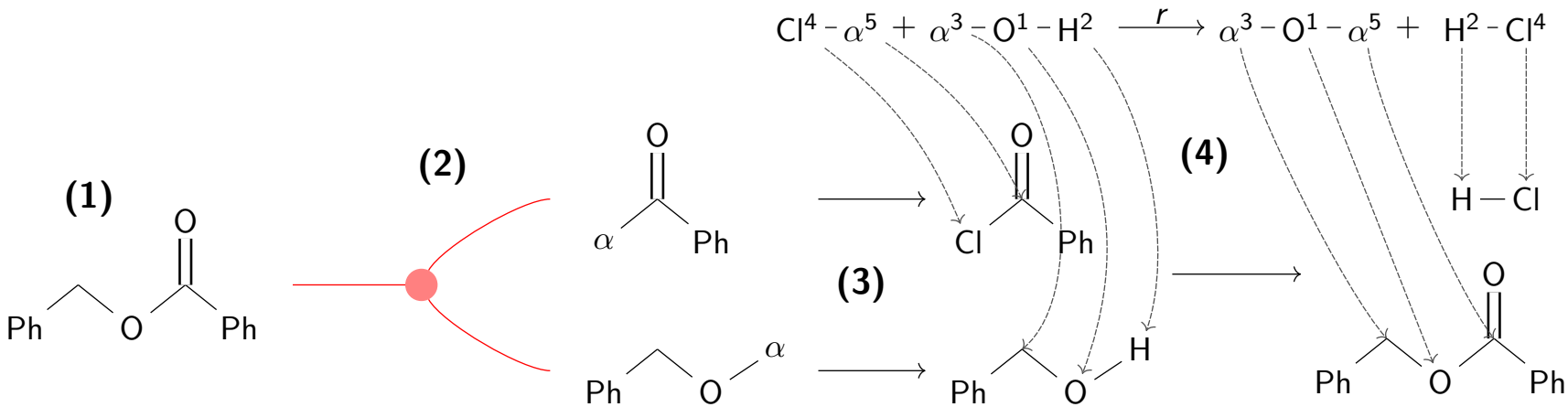
Retrosynthetic analysis

(1) Start with the target molecule(s)

(2) Cut the target along some bond, creating *synthons*

(3) Search for *synthetic equivalents*

(4) Search for a reaction whose reactants contain the synthetic equivalents, and whose products contain the target



Retrosynthetic analysis

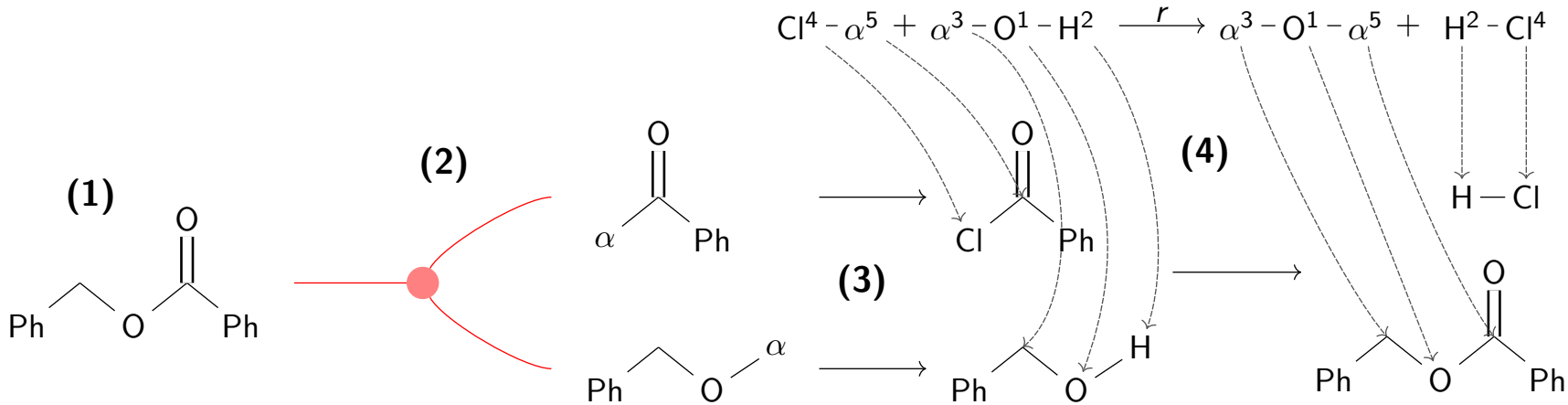
(1) Start with the target molecule(s)

(2) Cut the target along some bond, creating *synthons*

(3) Search for *synthetic equivalents*

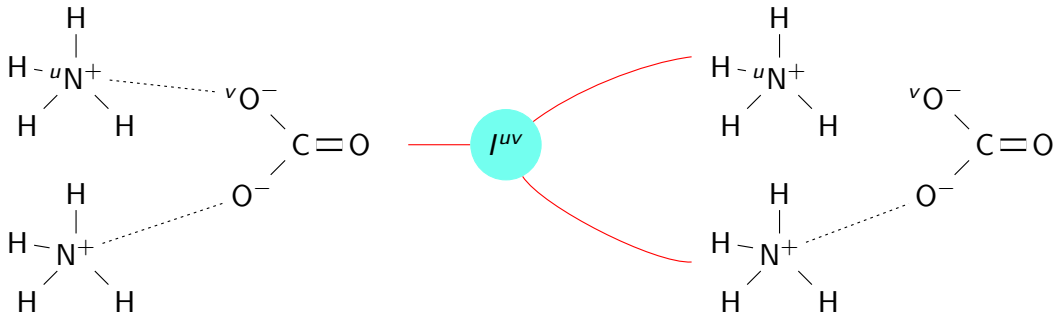
(4) Search for a reaction whose reactants contain the synthetic equivalents, and whose products contain the target

(5) Check whether the synthetic equivalents are known molecules: if yes, terminate, if no, return to **(1)** taking them as the target



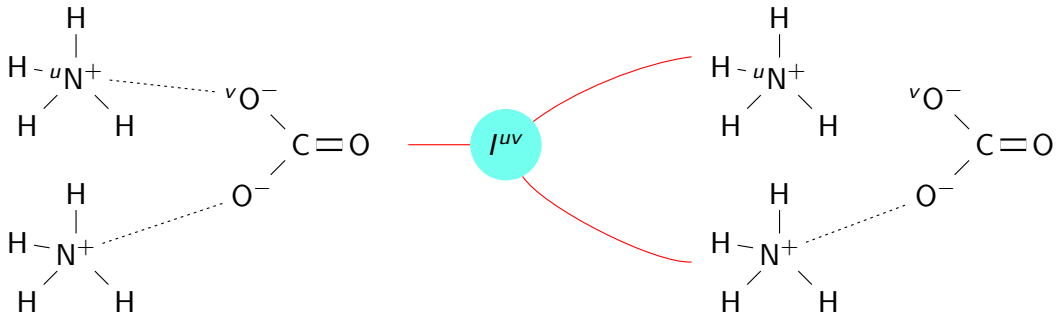
Disconnection rules

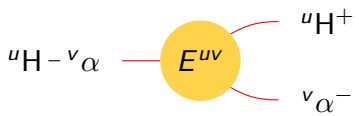
Motto: Chemical reactions are movements of electrons



Disconnection rules

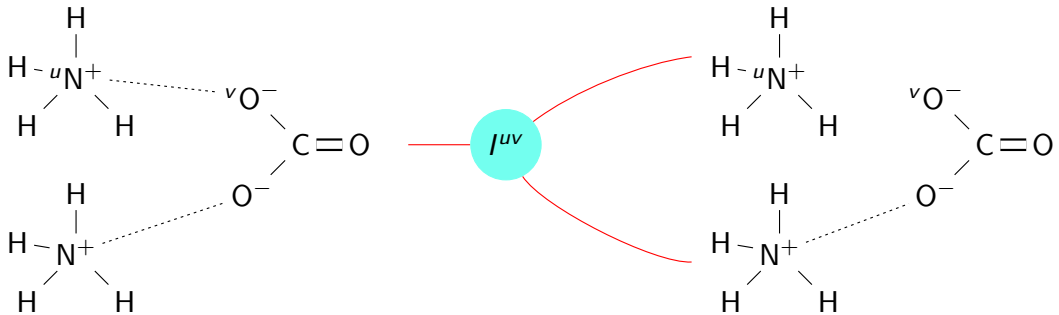
Motto: Chemical reactions are movements of electrons

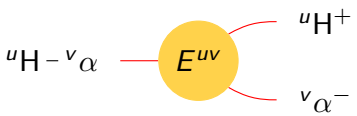


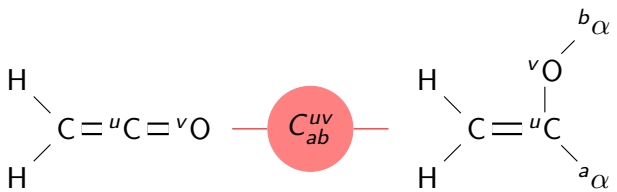


Disconnection rules

Motto: Chemical reactions are movements of electrons

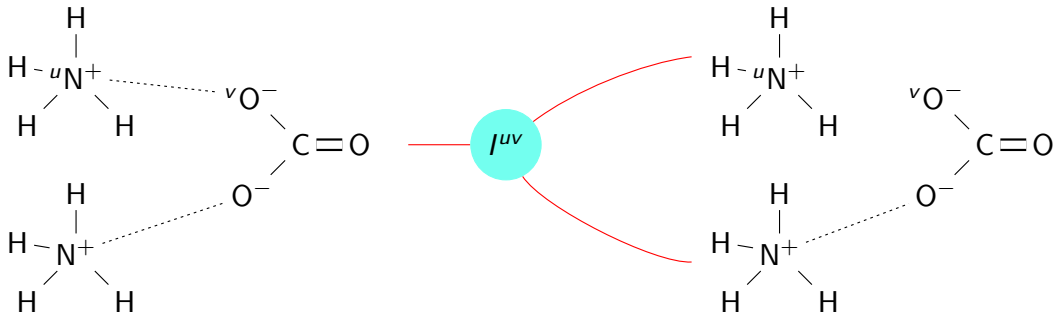


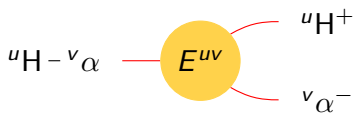


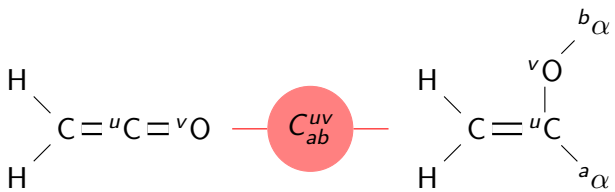


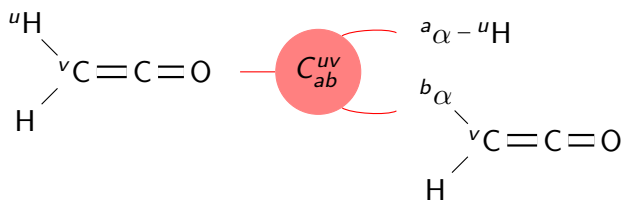
Disconnection rules

Motto: Chemical reactions are movements of electrons



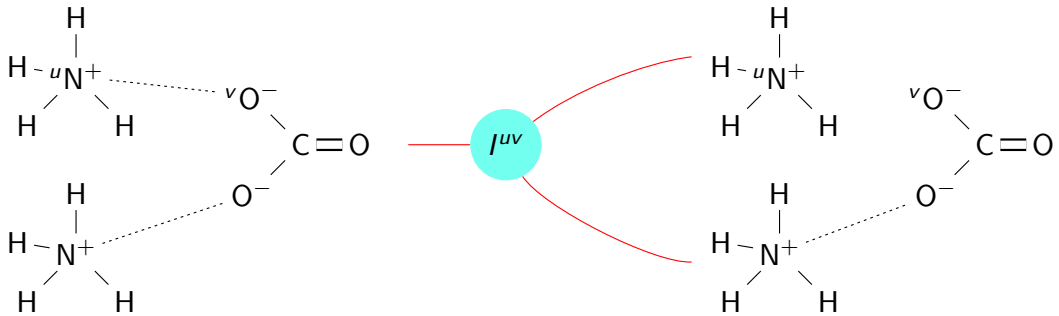


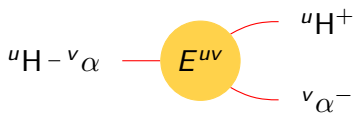


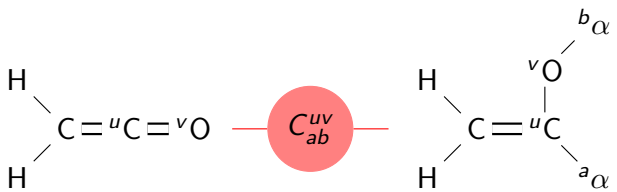


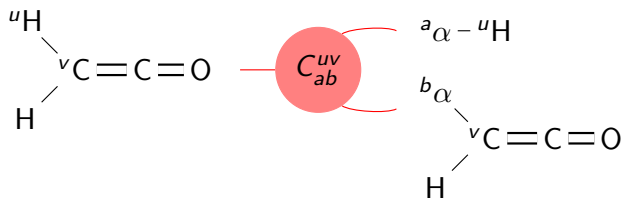
Disconnection rules

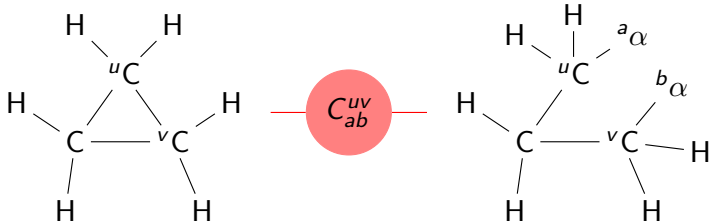
Motto: Chemical reactions are movements of electrons





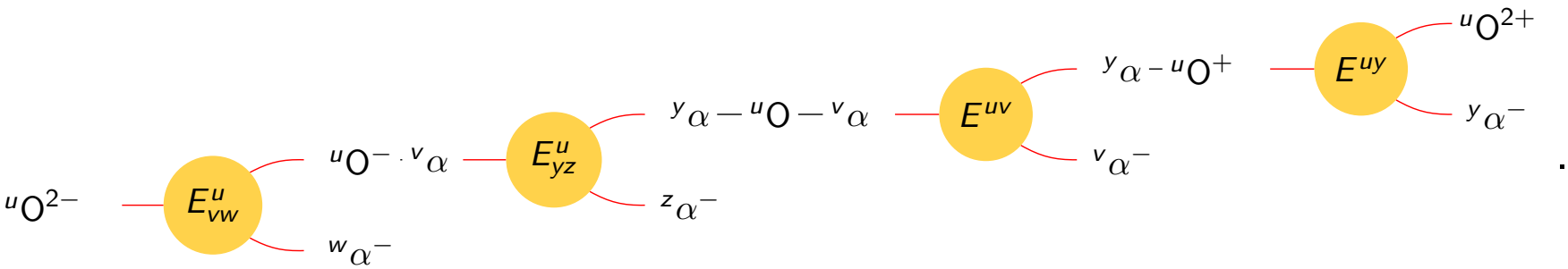






Disconnection rules

Motto: Chemical reactions are movements of electrons



$\{u\} \longrightarrow \{u, y, z\}$

$\{u, v\} \longrightarrow \{u, v\}$

$\{u, v\} \longrightarrow \{u,v,y,z\}$

Disconnection category

We define the set of terms with types as follows

- $\mathrm{id} : A \to A$,
- if $u \in \mathtt{Chem}(A)$, let $S^u : A \to A$,
- if $v \notin V_A \setminus \{u\}$, let $R^{u \mapsto v} : A \to A(u \mapsto v)$,
- $d^{uv}_{ab} : A \to d^{uv}_{ab}(A)$ and $\bar{d}^{uv}_{ab} : d^{uv}_{ab}(A) \to A$,
- if $\mathtt{t} : A \to B$ and $\mathtt{s} : B \to C$, then $\mathtt{t}; \mathtt{s} : A \to C$,
- if $\mathtt{t} : A \to B$ and $\mathtt{s} : C \to D$, then $\mathtt{t} + \mathtt{s} : A + C \to B + D$.

Disconnection category

We define the set of terms with types as follows

- $\mathrm{id} : A \to A$,
- if $u \in \mathtt{Chem}(A)$, let $S^u : A \to A$,
- if $v \notin V_A \setminus \{u\}$, let $R^{u \mapsto v} : A \to A(u \mapsto v)$,
- $d^{uv}_{ab} : A \to d^{uv}_{ab}(A)$ and $\bar{d}^{uv}_{ab} : d^{uv}_{ab}(A) \to A$,
- if $\mathtt{t} : A \to B$ and $\mathtt{s} : B \to C$, then $\mathtt{t}; \mathtt{s} : A \to C$,
- if $\mathtt{t} : A \to B$ and $\mathtt{s} : C \to D$, then $\mathtt{t} + \mathtt{s} : A + C \to B + D$.

The *disconnection category* **Disc** has valence complete chemical graphs as objects, and the terms (up to some equations) as morphisms.

From disconnections to reactions

Define the translation $R : \mathbf{Disc} \to \mathbf{React}$ by

- $R(E^u_{ab}) := (\{u\}, \{u, a, b\})$,
- $R(E^{uv}) := (\{u, v\}, \{u, v\})$,
- $R(I^{uv}) := (\{u\}, \{u\})$,
- $R(C^{uv}_{ab}) := (\{u\}, \{u, a, b\})$,
- $R(\mathrm{id}_A) := (\varnothing, \varnothing)$,
- $R(S^u) := (\{u\}, \{u\})$,
- $R(R^{u \mapsto v}) := (\varnothing, \varnothing)$,
- $R\left(\bar{d}^{uv}_{ab}\right) := \overline{R\left(d^{uv}_{ab}\right)}$,
- $R(\mathtt{t}; \mathtt{s}) := R(\mathtt{t}); R(\mathtt{s})$,
- $R(\mathtt{t} + \mathtt{s}) := R(\mathtt{t}) + R(\mathtt{s})$.

From disconnections to reactions

Proposition (Soundness)

$R : \mathbf{Disc} \to \mathbf{React}$ *is a monoidal dagger functor.*

From disconnections to reactions

Proposition (Soundness)

$R : \mathbf{Disc} \to \mathbf{React}$ *is a monoidal dagger functor.*

Theorem (Completeness)

$R : \mathbf{Disc} \to \mathbf{React}$ *is faithful.*

From disconnections to reactions

Proposition (Soundness)

$R : \mathbf{Disc} \to \mathbf{React}$ *is a monoidal dagger functor.*

Theorem (Completeness)

$R : \mathbf{Disc} \to \mathbf{React}$ *is faithful.*

Theorem

The composite functor

$$\mathbf{Disc} \xrightarrow{R} \mathbf{React} \xrightarrow{[-]} \mathbf{React}_{\sim}$$

is an opfibration.

From disconnections to reactions

Proposition (Soundness)

$R : \mathbf{Disc} \to \mathbf{React}$ *is a monoidal dagger functor.*

Theorem (Completeness)

$R : \mathbf{Disc} \to \mathbf{React}$ *is faithful.*

Theorem

The composite functor

$$\mathbf{Disc} \xrightarrow{R} \mathbf{React} \xrightarrow{[-]} \mathbf{React}_{\sim}$$

is an opfibration.

Corollary (Universality)

Given a reaction $r : A \to B$, there is a sequence of (dis)connection rules $d : A \to B'$ such that $B \simeq B'$.

Conclusion

- Connected reaction representation with retrosynthetic (dis)connection rules

Conclusion

- Connected reaction representation with retrosynthetic (dis)connection rules
- Showed that all (formal) reactions can be decomposed into (dis)connection rules

Conclusion

- Connected reaction representation with retrosynthetic (dis)connection rules
- Showed that all (formal) reactions can be decomposed into (dis)connection rules

Future work:

Conclusion

- Connected reaction representation with retrosynthetic (dis)connection rules
- Showed that all (formal) reactions can be decomposed into (dis)connection rules

Future work:

- What about reactions that are not valence complete?

Conclusion

- Connected reaction representation with retrosynthetic (dis)connection rules
- Showed that all (formal) reactions can be decomposed into (dis)connection rules

Future work:

- What about reactions that are not valence complete?
- Incorporate chirality into the model

Conclusion

- Connected reaction representation with retrosynthetic (dis)connection rules
- Showed that all (formal) reactions can be decomposed into (dis)connection rules

Future work:

- What about reactions that are not valence complete?
- Incorporate chirality into the model
- Implement the algorithm that turns reactions into disconnection rules in normal form

References

- Andersen, Flamm, Merkle, Stadler. *Inferring chemical reaction patterns using rule composition in graph grammars*. Journal of Systems Chemistry. 2013.
- Clayden, Greeves, Warren. *Organic chemistry*. Oxford University Press. 2012.
- Gale, Lobski, Zanasi. *A categorical approach to synthetic chemistry*. International Colloquium on Theoretical Aspects of Computing. 2023.
- Habel, Plump. *$\mathcal{M},\mathcal{N}$-Adhesive Transformation Systems*. International Conference on Graph Transformation. 2012.
- Warren, Wyatt. *Organic synthesis: the disconnection approach*. Wiley. 2008.

Thank you for your attention!