SYSTEMATIC METHODS FOR SOLVENT DESIGN: TOWARDS BETTER REACTIVE PROCESSES

Eirini Siougkrou

National Technical University of Athens

KT Consortium annual meeting, DTU
8 June 2017

Outline

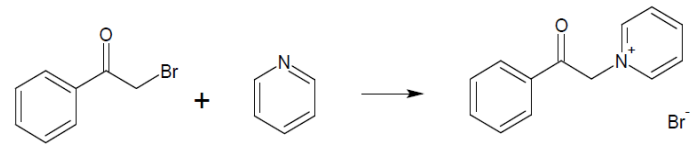
- Systematic methods for solvent design
  - A methodology for the integrated design of a gas-expanded liquid and reactive system
  - *Ab-Initio* Computer-Aided Molecular Design for the Identification of Optimal Solvents for Reactions
- Semantics and Process Systems Engineering

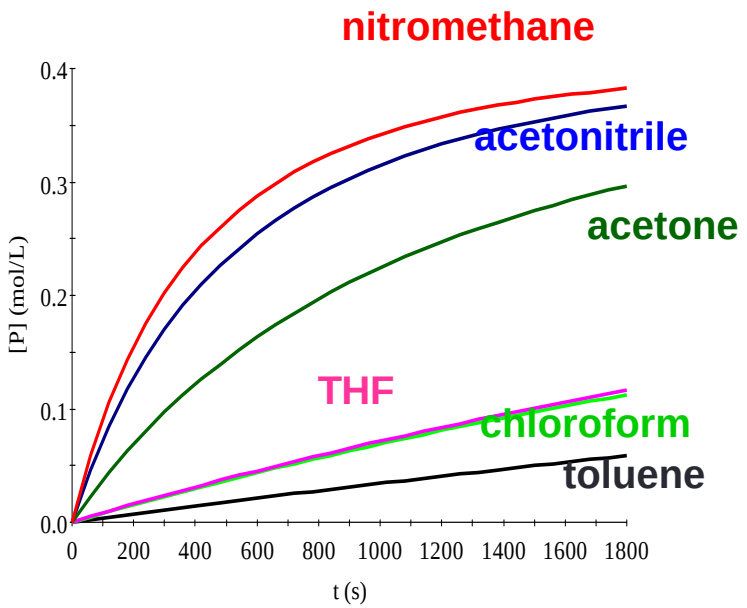
Outline

- Systematic methods for solvent design
  - A methodology for the integrated design of a gas-expanded liquid and reactive system
  - *Ab-Initio* Computer-Aided Molecular Design for the Identification of Optimal Solvents for Reactions
- Semantics and Process Systems Engineering

Solvents and chemical reactions

- The importance of solvents in chemical reactions is well known:
  - bring reactants together
  - temperature control
  - selectivity/reaction rate
  - transport
  - separation
- Reaction rate constants can vary by several orders of magnitude from solvent to solvent.





The concentration of the product for different solvents for the Menschutkin reaction of phenacyl bromide and pyridine[1]:

H. Struebing et al., 2013

Solvent optimisation very important
Need for alternative “green” solvents

Objectives

- The development of novel methodologies for the design of solvents for chemical reactions

- Two aspects considered:

  Design of solvent mixture
  - CO_2-expanded solvents
  - experimental data needed
  - integration in process design

  Ab-initio solvent design
  - no experimental data
  - generic QM-CAMD methodology

Gas-Expanded Liquids (GXLs)

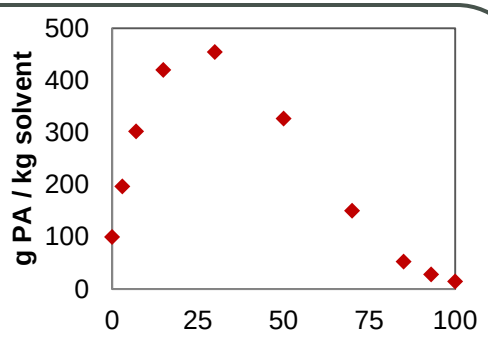
GXLs[1] are **mixed solvents** composed of:

organic solvent + compressible gas (usually CO_2)

- Why mixed solvents?

- tuneable properties
  - combination of desirable properties of co-solvents
  - elimination of undesirable properties of co-solvents

Solubility of paracetamol in water+acetone mixture, at T = 23 °C.



R.A. Granberg, A.C.Rasmuson, 2000

P.G.Jessop, B.Subramanian, 2007

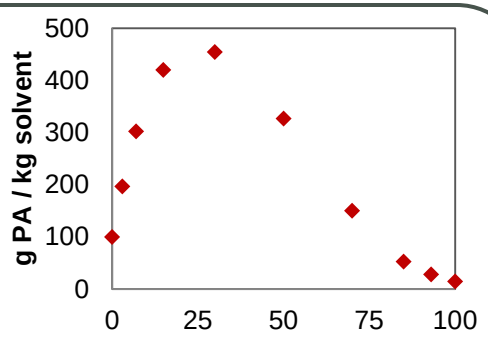
Gas-Expanded Liquids (GXLs)

GXLs[1] are **mixed solvents** composed of:

organic solvent + compressible gas (usually CO_2)

- Why mixed solvents?
  - tuneable properties

Solubility of paracetamol in water+acetone mixture, at T = 23 °C.



R.A. Granberg, A.C.Rasmuson, 2000

- Why GXLs?
  - recovery and recycle of both the organic solvent and CO_2 through depressurisation
  - moderate operating pressures
  - enhanced transport rates and reaction rates
  - reduced environmental impact

✓ GXLs meet process and environmental requirements

P.G.Jessop, B.Subramanian, 2007

Objectives

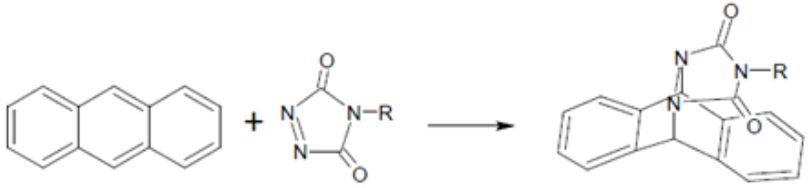
To **develop a methodology** for the integrated design of a reactive system including a solvent mixture.

Given a reaction and a production rate, find the optimal

- CO_2-expanded solvent
- equipment size
- operating conditions

that minimise the total cost of the process.

Case study: the Diels-Alder reaction of anthracene with PTAD.



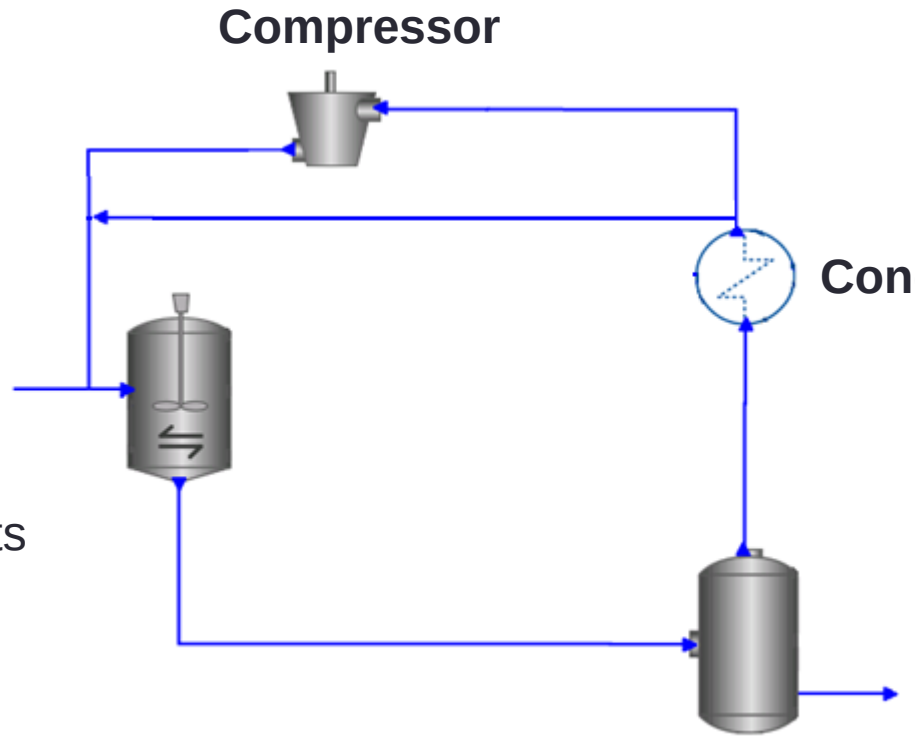
Challenges: high pressure

presence of solids

effect of solvent composition on reaction rate

Our Model

- For the process we consider a **CSTR**, a **separator** (evaporator), a **condenser** and a **compressor**.



CSTR
- T = 40 ^{o}C
- Neglect heat effects

Separator
- treated as black box
- solvent evaporation only
- 100% recovery

The model is implemented in **gPROMS**[1].

1. Process Systems Enterprise, gPROMS, 1997-2009

Our Model

- **The model can be divided in five sub-models:**

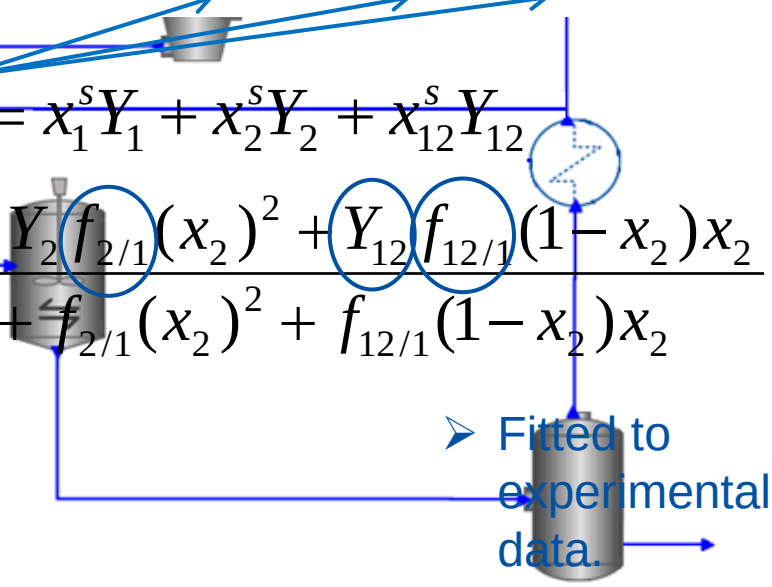
- Reaction rate constant model: **Solvatochromic equation**[1]
- Composition dependence of solvent properties[2]
- Phase Equilibrium: **Group-contribution VTPR EoS**[3]
- Process mass balances
- Cost model

- Organic co-solvents: **acetonitrile**, **acetone**, **methanol**.

$$k = k_0 + s\pi^* + a\alpha + b\beta$$

$$Y = x_1^S Y_1 + x_2^S Y_2 + x_{12}^S Y_{12}$$

$$Y = \frac{Y_1(1-x_2)^2 + Y_2 f_{2/1}(x_2)^2 + Y_{12} f_{12/1}(1-x_2)x_2}{(1-x_2)^2 + f_{2/1}(x_2)^2 + f_{12/1}(1-x_2)x_2}$$

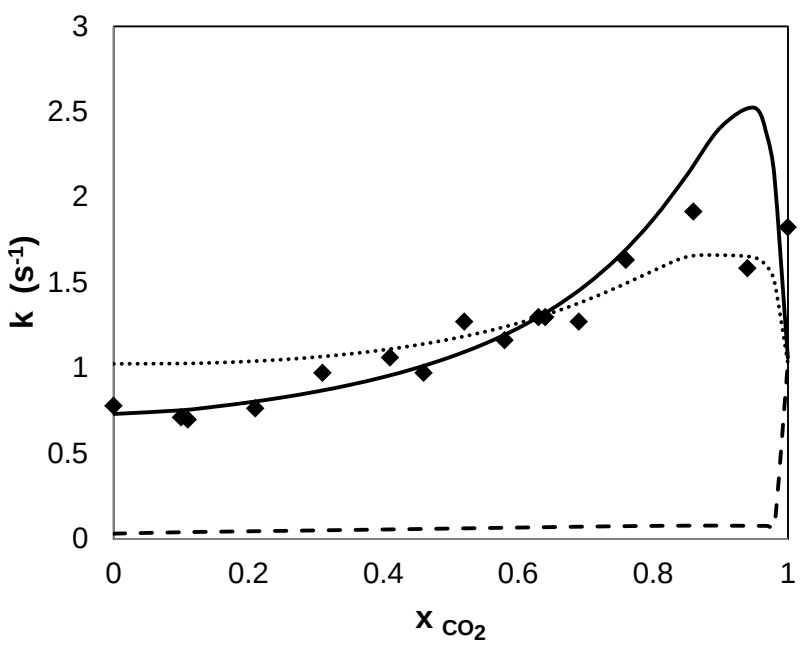


- Fitted to experimental data.

1. J.W. Ford, J.Lu, C.L.Liotta, C.A.Eckert, 2008
2. C.Ràfols, M.Rosés, E.Bosch, 1997
3. J.Ahlers, T.Yamaguchi, J.Gmehling, 2004

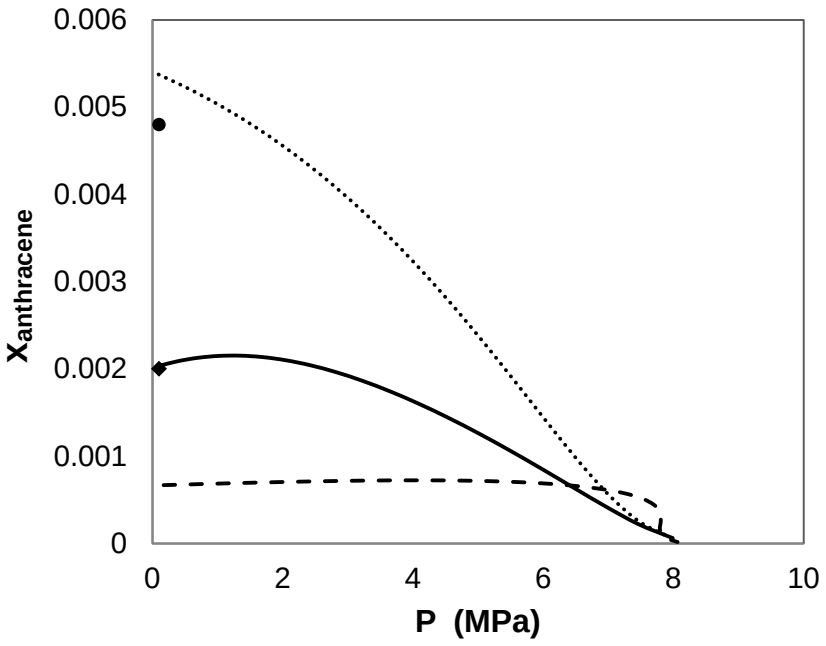
E. Siougkrou, A. Galindo, C.S. Adjiman, 2011
E. Siougkrou, A. Galindo, C.S. Adjiman, 2014

Reaction Rate Constant Solubility



- The reaction rate constant increases with increasing x_{CO_2} in all mixed solvents.

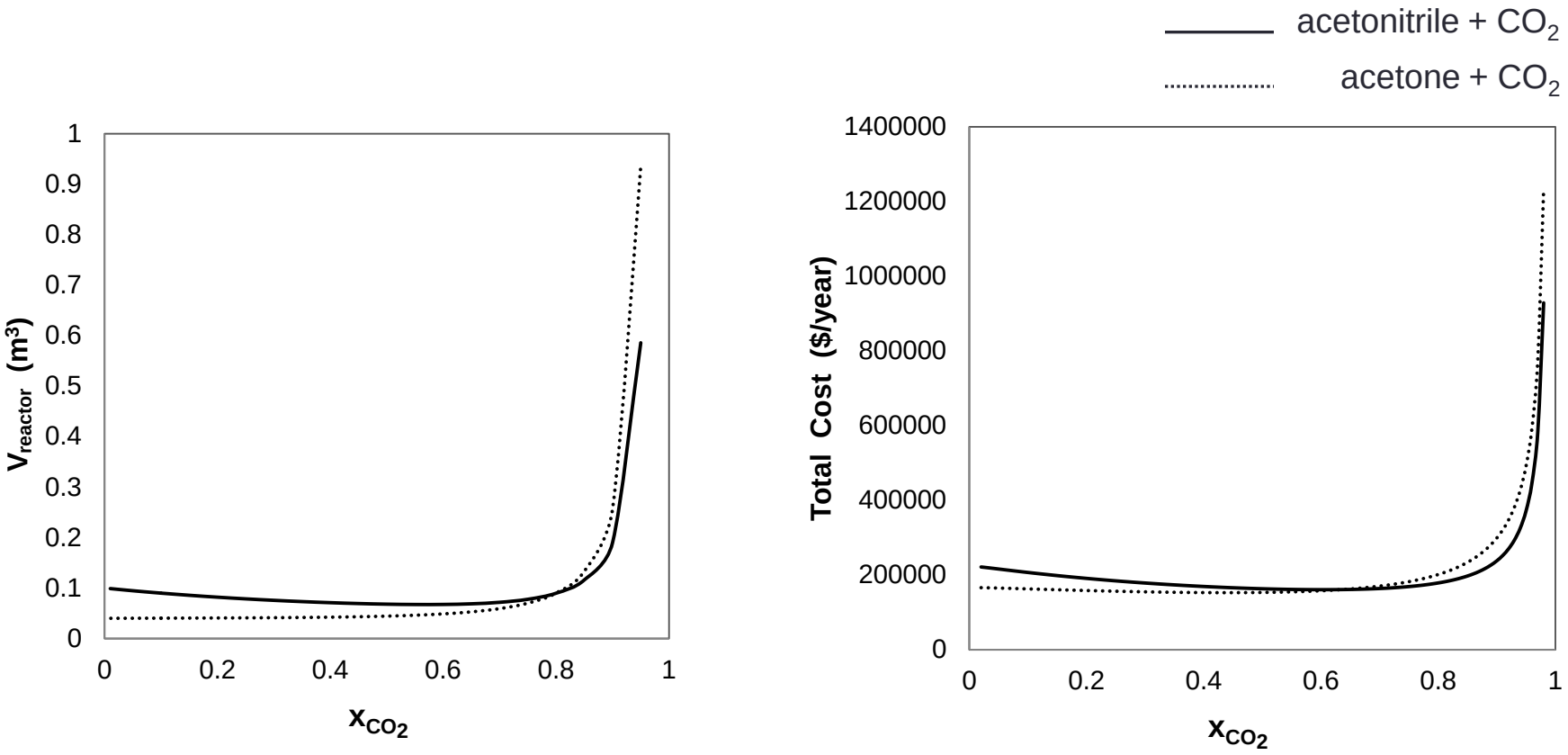
J.W. Ford, J.Lu, C.L.Liotta, C.A.Eckert, 2008
E.A.Cepeda, M.Diaz, 1996
L.N. Petrova, 1974



- The solubility of anthracene decreases with increasing x_{CO_2} in all mixed solvents.

E. Siougkrou, A. Galindo, C.S. Adjiman, 2014

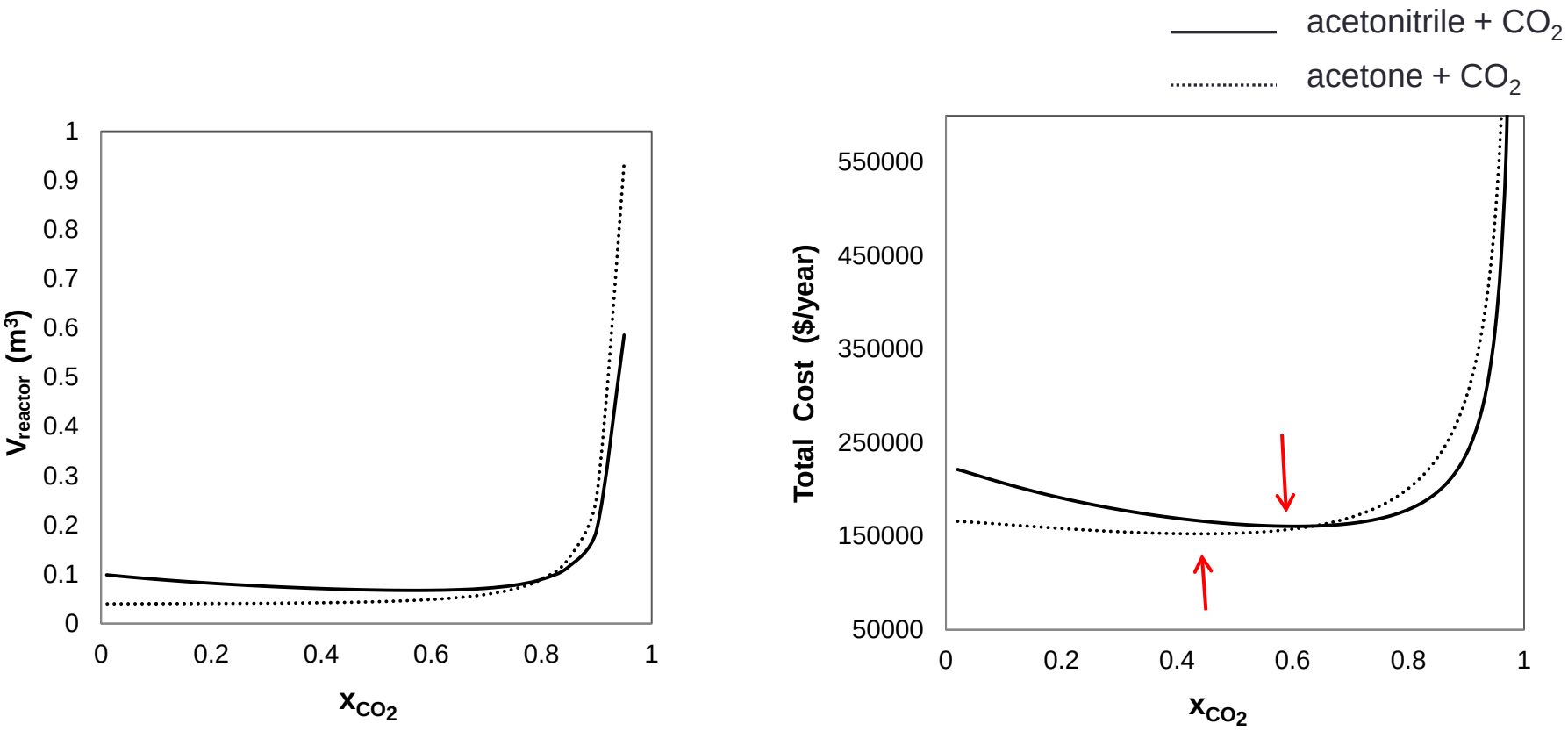
Volume of the CSTR and Cost of the Process



- Methanol not shown as it requires very large volumes.
- Both acetonitrile and acetone seem to be suitable co-solvents.
- Optimum $\mathbf{x_{CO_2} = 0.45}$ in acetone.
- Taking into account environmental criteria, the **optimum x_{CO_2} is around 0.7.**

E. Siougkrou, A. Galindo, C.S. Adjiman, 2014

Volume of the CSTR and Cost of the Process



- Methanol not shown as it requires very large volumes.
- Both acetonitrile and acetone seem to be suitable co-solvents.
- Optimum $\mathbf{x_{CO_2} = 0.45}$ in acetone.
- Taking into account environmental criteria, the **optimum x_{CO_2} is around 0.7.**

E. Siougkrou, A. Galindo, C.S. Adjiman, 2014

Outline

- A methodology for the integrated design of a gas-expanded liquid and reactive system

- *Ab-Initio* Computer-Aided Molecular Design for the Identification of Optimal Solvents for Reactions

Objectives

To develop a methodology for the design of the optimal solvent that maximises the rate constant of a given reaction.

- Enables solution of more complex problems to be tackled
  - e.g., selectivity maximisation

- Requirements of the methodology:
  - Consider a large number of candidate solvents and take into account multiple constraints & objectives ⇒
    Computer-Aided Molecular Design (CAMD)
  - Based on reliable prediction of kinetics ⇒
    Quantum Mechanics (QM) + **Group Contribution** (GC)
  - Computational efficiency ⇒ **limit number of QM calculations**

Solvent design problem

A Computer-Aided Molecular Design optimisation formulation

$$\max_{\xi,\mathbf{n},\mathbf{y}} f(\xi)$$

$$h_1(\xi,\mathbf{n},\mathbf{y}) = 0$$
$$g_1(\xi,\mathbf{n},\mathbf{y}) \le 0$$

← structure-property constraints

$$h_2(\mathbf{n},\mathbf{y}) = 0$$
$$g_2(\mathbf{n},\mathbf{y}) \le 0$$

← chemical feasibility and molecular complexity constraints

$$d(\xi,\mathbf{n},\mathbf{y}) \le 0$$ design constraints

$$\xi \in \Re^{m}$$ physical properties

$$\mathbf{n} \in \Re^{q}$$ number of groups in molecule

$$\mathbf{y}_i \in \{0,1\}^{u} \quad i = 1,\ldots,q$$ binary variables

Solvent design problem

However, the consideration of reaction kinetics brings significant challenges

$$\max_{\xi,\mathbf{n},\mathbf{y}} f(\xi)$$

$$h_1(\xi,\mathbf{n},\mathbf{y}) = 0$$
$$g_1(\xi,\mathbf{n},\mathbf{y}) \le 0$$
$$h_2(\mathbf{n},\mathbf{y}) = 0$$
$$g_2(\mathbf{n},\mathbf{y}) \le 0$$
$$d(\xi,\mathbf{n},\mathbf{y}) \le 0$$
$$\xi \in \mathrm{R}^{\mathrm{m}}$$
$$\mathbf{n} \in \mathrm{R}^{\mathrm{q}}$$
$$\mathbf{y}_{\mathrm{i}} \in \{0,1\}^{u} \quad i = 1,\ldots,\mathrm{q}$$

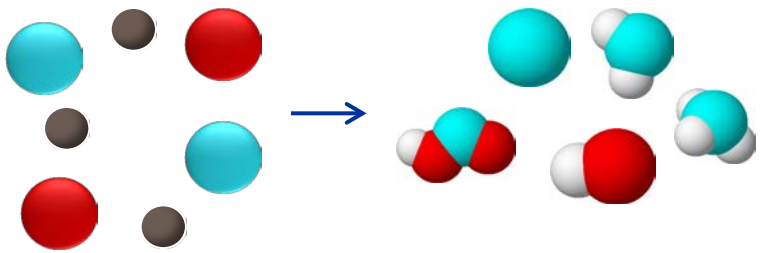
reaction rate constant

structure-property constraints

Property prediction for solvent design

Atoms

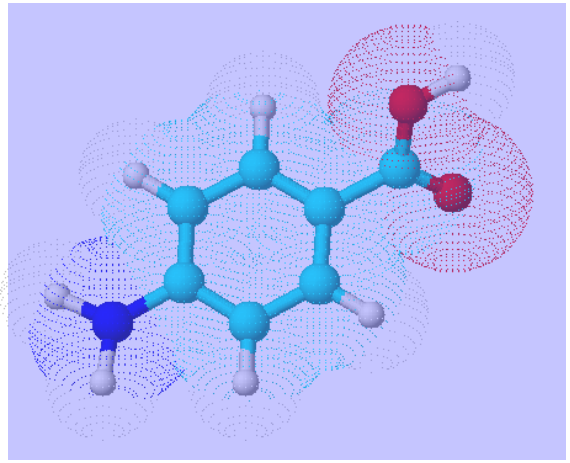
Group Contribution methods[1]



A
B
ε
γ
n_D
φ
ψ

solvent at macroscale

Continuum Solvation[2]
Quantum Mechanics



k

solute at electronic scale
solvent at macroscale

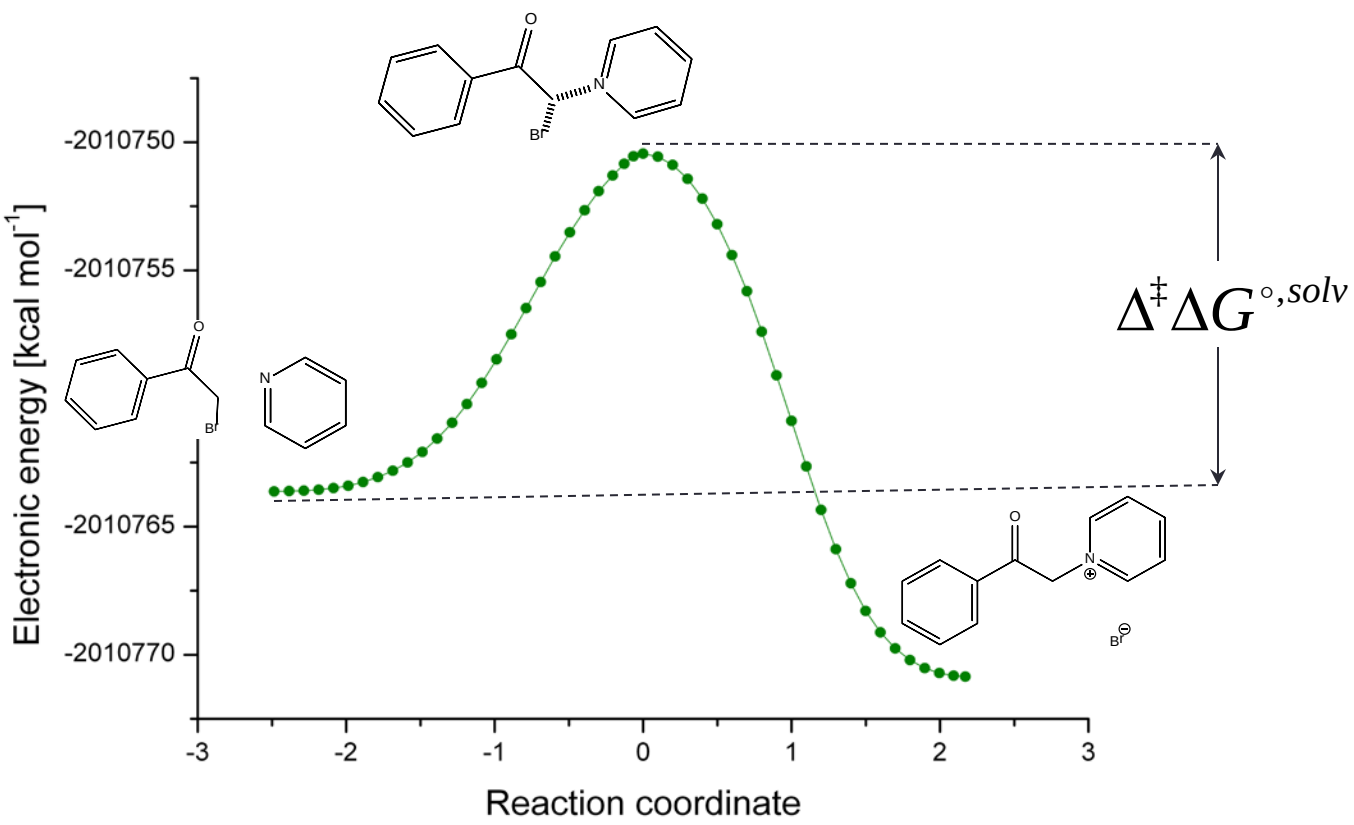
- described with bulk properties

1. T. J. Sheldon et al., Fluid Phase Equilibria 231 (2005) 27-37
2. A. V. Marenich et al., J.Phys. Chem. B 113 (2009) 6378-6396

Reaction rate constant from Quantum Mechanics

- Rate constant from **Conventional Transition State Theory** (CTST):

$$R_1 + R_2 \leftrightarrow TS \rightarrow P$$



Reaction rate constant from Quantum Mechanics Free energy of solvation

- Rate constant from **Conventional Transition State Theory** (CTST) using the SMD solvation model:

$$R_1 + R_2 \leftrightarrow TS \rightarrow P$$

$$k_{\mathrm{TST}} = f\left(\Delta^{\ddagger}\Delta G^{\circ,solv}(\varepsilon, A, B, \gamma, n_D, \varphi, \psi)\right)$$

activation free energy of solvation

$$\Delta G_i^{o,solv}(\xi_S) = \min_{\mathbf{r}_i} \Delta E_{ENP,i}(\mathbf{r}_i; \varepsilon, A) + G_{CDS,i}\left(\mathbf{r}_i^*(\varepsilon, A), A, B, \gamma, n_D, \varphi, \psi\right)$$

electrostatic contribution

non-electrostatic contribution

$$\mathbf{r}_i^*(\varepsilon, A) = \arg\min_{\mathbf{r}_i} \Delta E_{ENP,i}(\mathbf{r}_i; \varepsilon, A)$$

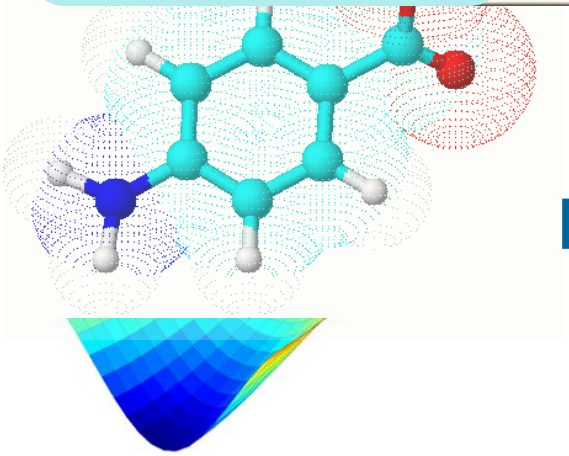
Bilevel problem with embedded QM calculations

QM-CAMD solvent design methodology

Step 1.
Define the initial solvent set

Step 2.
lculate rate
ant in specific
ent(s) by QM

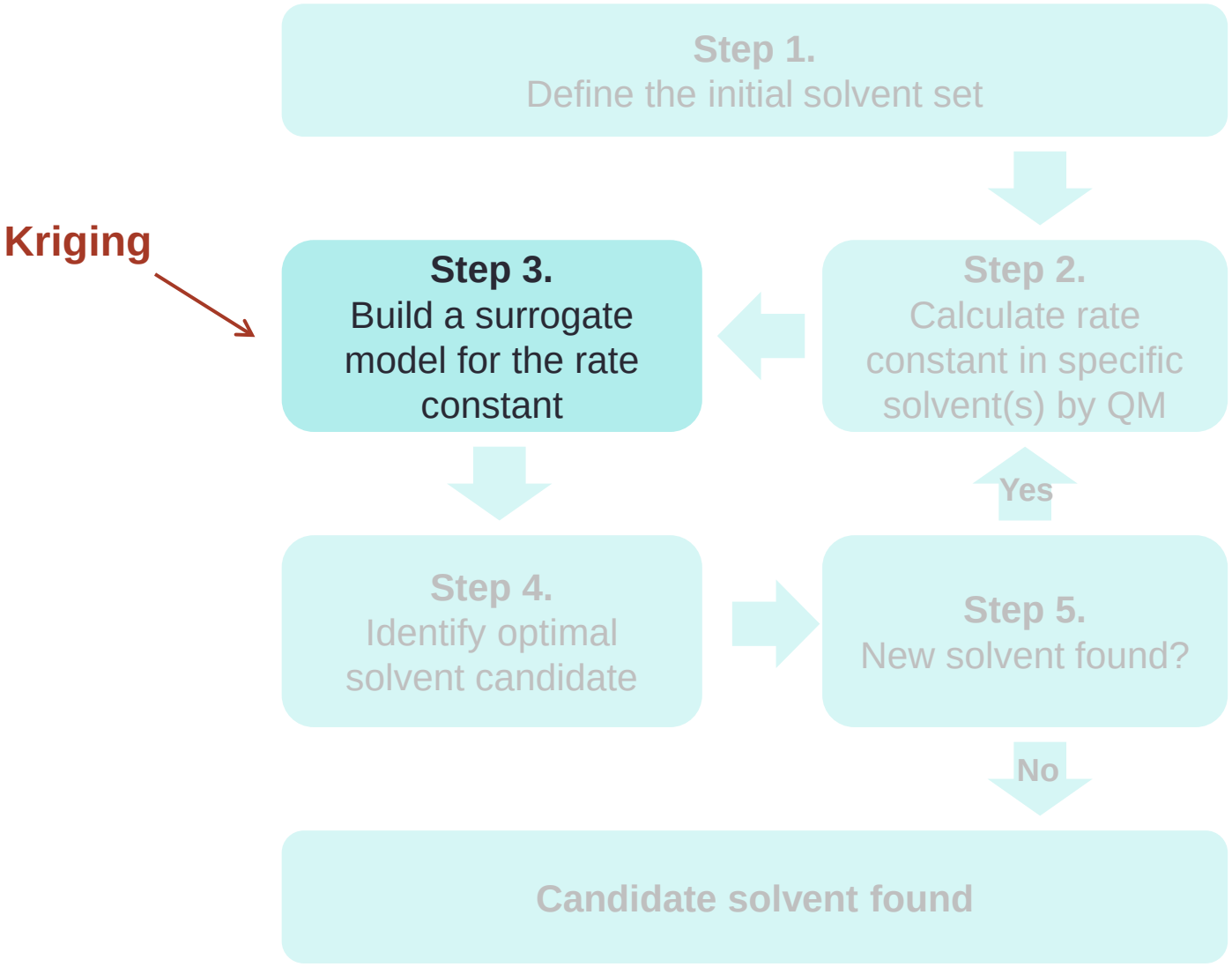
Step 3.
Build a surrogate model for the rate constant



Candidate solvent found

H. Struebing et al., 2013

QM-CAMD-Kriging solvent design methodology



E. Siougkrou, A. Galindo, C.S. Adjiman, in preparation

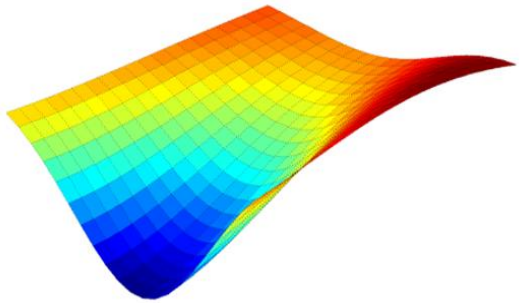
Choice of surrogate model

- Use a “physical” model, e.g. a linear free energy relation
  - solvatochromic equation – a simple model
  - approach successfully demonstrated on a S_N2 reaction[1]
  - inherent mismatch between detailed and surrogate models makes identification of best solvent challenging

- Use a response-surface methodology
  - ensure convergence between detailed and surrogate models
  - … at the cost of increased computational complexity

1. **H. Struebing et al.,, 2013**

Building a Kriging response surface



- **What makes Kriging special**
  - ✓ Exact extrapolator with a statistical interpretation
  - ✓ Correlation between two random variables

$$R_{ij} = \exp\left(-\sum_{l=1}^{7} \theta_l \left| \xi_{S,il} - \xi_{S,jl} \right|^{p_l} \right)$$

θ, p: adjustable parameters, $\xi_S^{\mathrm{T}} = (\varepsilon, \mathrm{A}, \mathrm{B}, \gamma, n_D, \phi, \psi)$ - solvent properties

- **Kriging predictor**

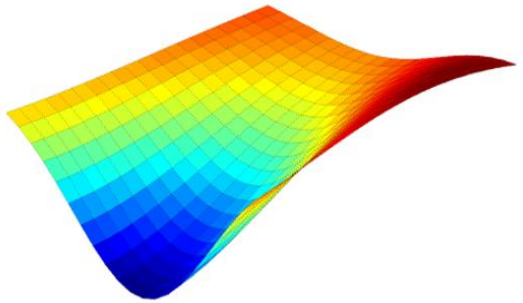
$$k_{KR}(\xi_S) = \hat{k}_{TST} + \mathbf{r}'\mathbf{R}^{-1}(\mathbf{k}_{TST} - \mathbf{1}\hat{k}_{TST})$$

$\hat{k}_{TST}$: mean value, $\mathbf{R}$: correlation matrix, $\mathbf{r}$: augmented correlation matrix

$\mathbf{k}_{TST}$: set of rate constants evaluated at Kriging by DFT+SMD

D. R. Jones, J. Global Opt. 21 (2001) 345-383

Building a Kriging response surface



- **What makes Kriging special**
  - ✓ Exact extrapolator with a statistical interpretation
  - ✓ Correlation between two random variables

$$R_{ij} = \exp\left(-\sum_{l=1}^{7} \theta_l \left| \xi_{S,il} - \xi_{S,jl} \right|^{p_l} \right)$$

θ, p: adjustable parameters, $\xi_S^{\mathrm{T}} = (\varepsilon, \mathrm{A}, \mathrm{B}, \gamma, n_D, \phi, \psi)$ - solvent properties

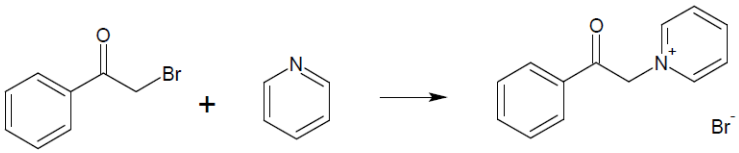
- **Kriging predictor**

$$k_{KR}(\xi_S) = \hat{k}_{TST} + \mathbf{r}'\mathbf{R}^{-1}(\mathbf{k}_{TST} - \mathbf{1}\hat{k}_{TST})$$

- $k_{KR} = k_{KR}(\varepsilon, A, B, \gamma, n_D, \varphi, \psi)$

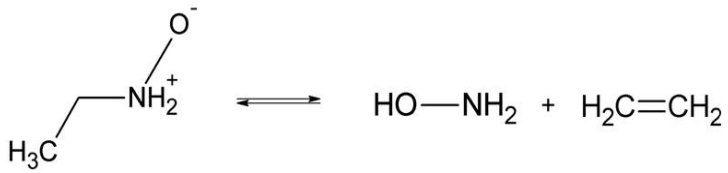
Application to two chemical reactions

Menschutkin reaction



- Best solvent designed: dichlromethanol
  - A **165% increase** in predicted rate constant
  - A **126% increase** over best solvent using linear surrogate model

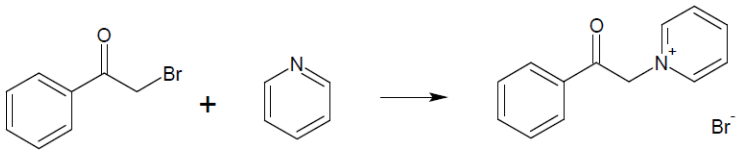
Cope elimination

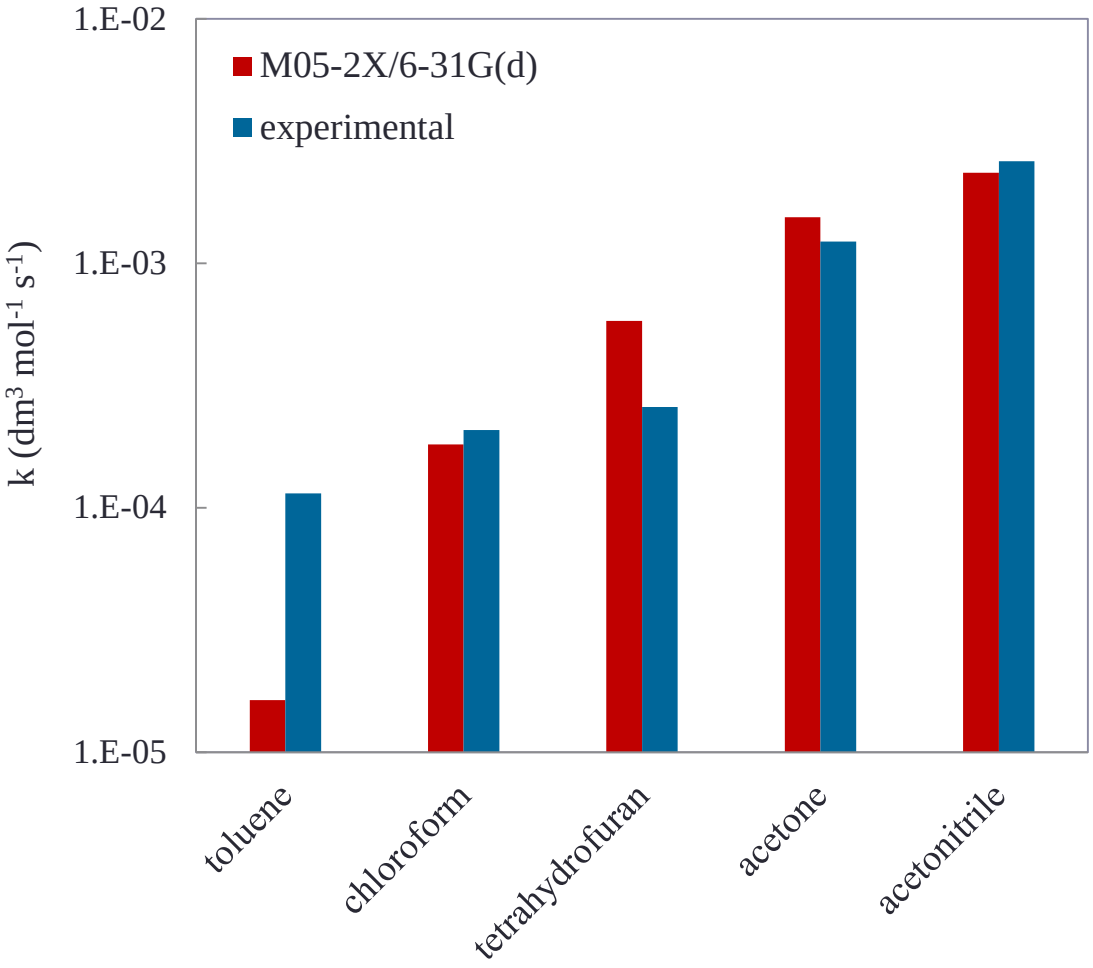


- Best solvent designed: methylpentane
  - A **326% increase** in predicted rate constant

E. Siougkrou, A. Galindo, C.S. Adjiman, in preparation

How reliable are the DFT+SMD calculations?

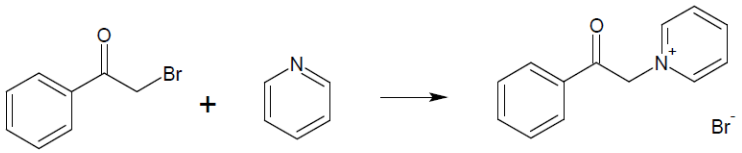


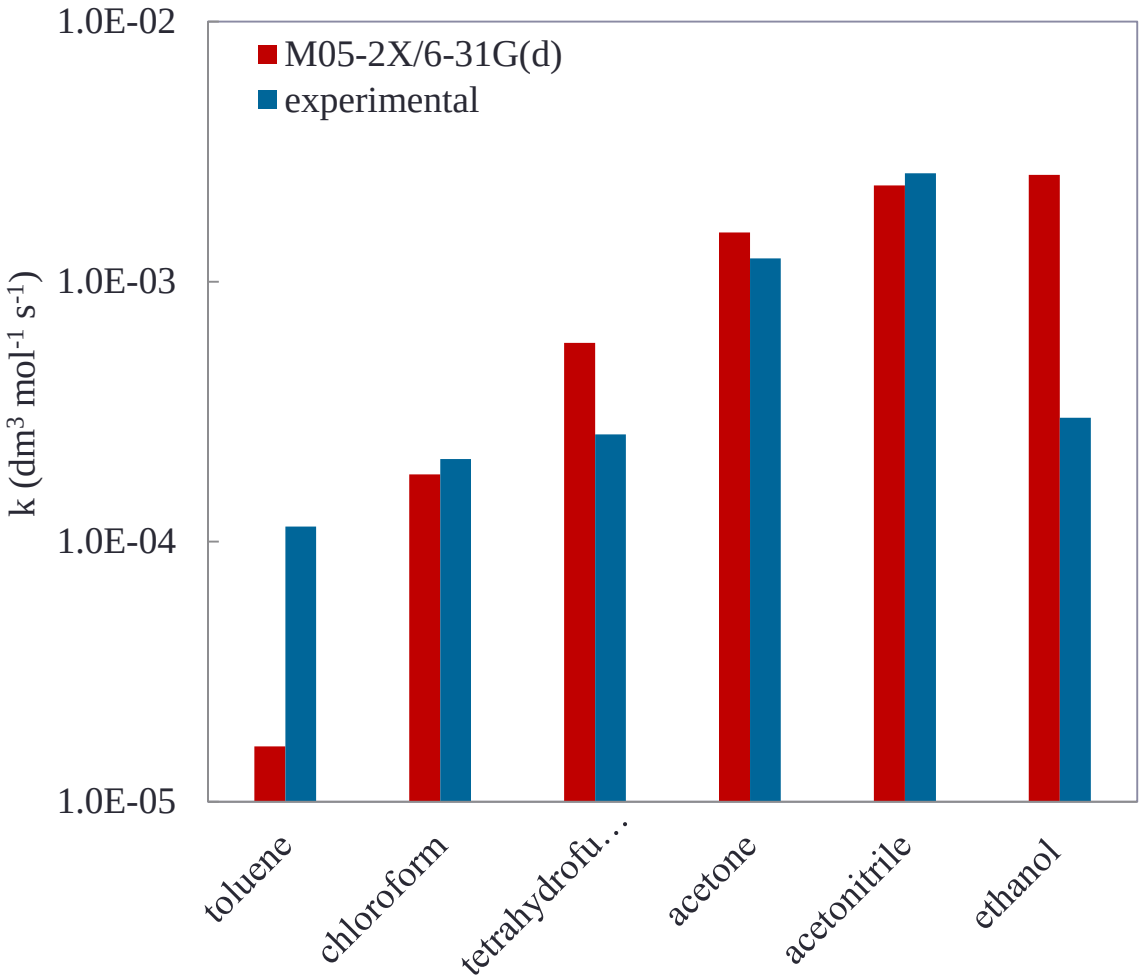


- Reaction studied at various levels of theories and basis sets
- QM calculations for rate constant in good agreement with experiments for aprotic solvents
- Match between experimental and computational rankings

H. Struebing et al. Nature Chem. 5 (2013) 952-957

How reliable are the DFT+SMD calculations?





- Not very good agreement in protic solvents
- Best aprotic solvent designed: $Cl(CH_2)_2NO_2$

  k_{TST} = 3.24 x 10^{-3} dm^3 mol^{-1} s^{-1}

  - A 35% increase

Summary & Perspectives

- A model formulation for the integrated design of a CO_2-expanded solvent and a reactive system has been developed
  - Need for new, more predictive models for solvent effects (focused on kinetics/rate constants)
- A new computational methodology for the design of solvents
  - Combination of QM, kriging and CAMD
- QM-CAMD methodology extended to account for solubility of reactants (H. Struebing, S. Oberhaimer, E. Siougkrou, C. Adjiman, A. Galindo, 2017)
- Investigation of the effect of the initial set of solvents for the QM-CAMD algorithm and optimal selection of the initial set (T. Oliyide, E. Siougkrou, C. Adjiman, in preparation)

Outline

- Systematic methods for solvent design
  - A methodology for the integrated design of a gas-expanded liquid and reactive system
  - *Ab-Initio* Computer-Aided Molecular Design for the Identification of Optimal Solvents for Reactions
- Semantics and Process Systems Engineering

Why semantics

- Natural language interpretation

’Would it be *profitable* to install a *bioethanol* plant in Greece?

cost/profit estimation
cost models for biorefinaries
symbiosis info?

product info
processing technologies

’What *product portfolios* would be possible?

processing technologies
possible value chains

’For which of these product portfolios are the *technologies* required existing?

existing technology models
experimental data

Why semantics

- Natural language interpretation
- Large amount of data
- Heterogeneous resources
  - mathematical models, flowsheets, experimental data, etc.
- High-throughput organisation and discovery of information
- Update of information without the need for redesign

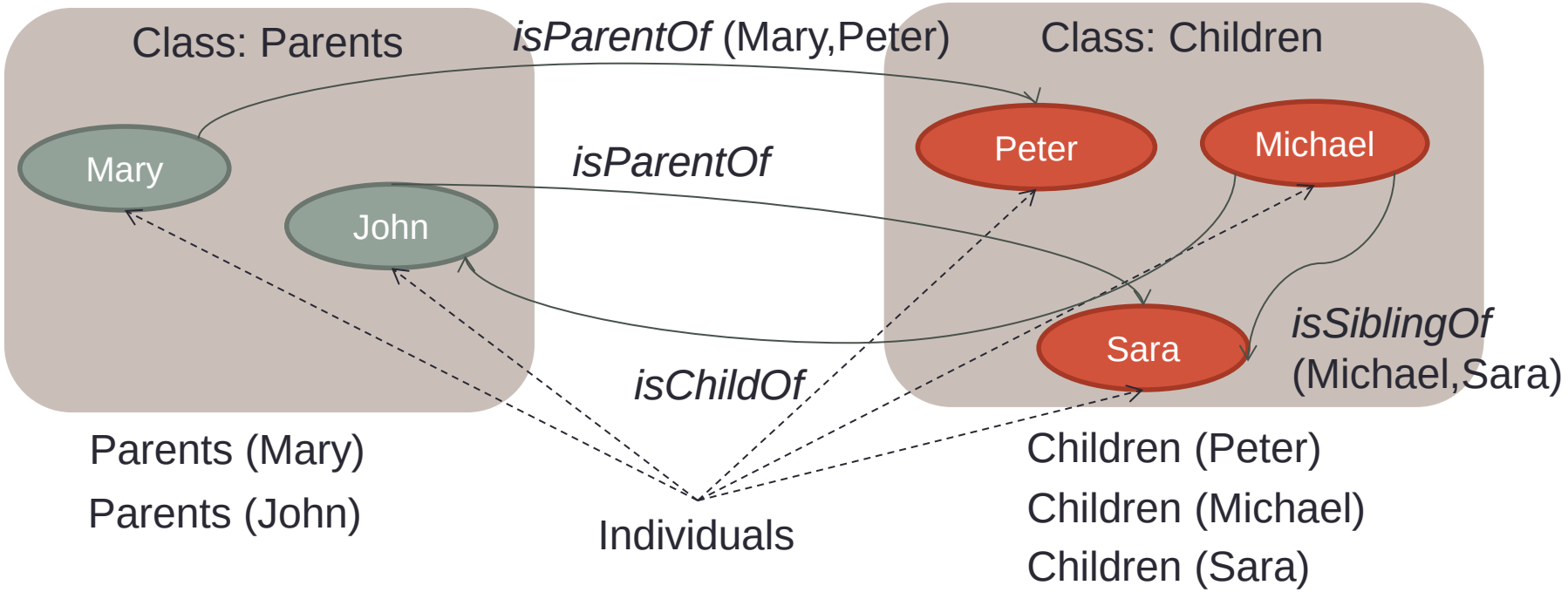
Ontology engineering

"*An ontology is an explicit, formal specification of a shared conceptualization. The term is borrowed from philosophy, where an Ontology is a systematic account of Existence. For Artificial Intelligence systems, what ''exists" is that which can be represented*."

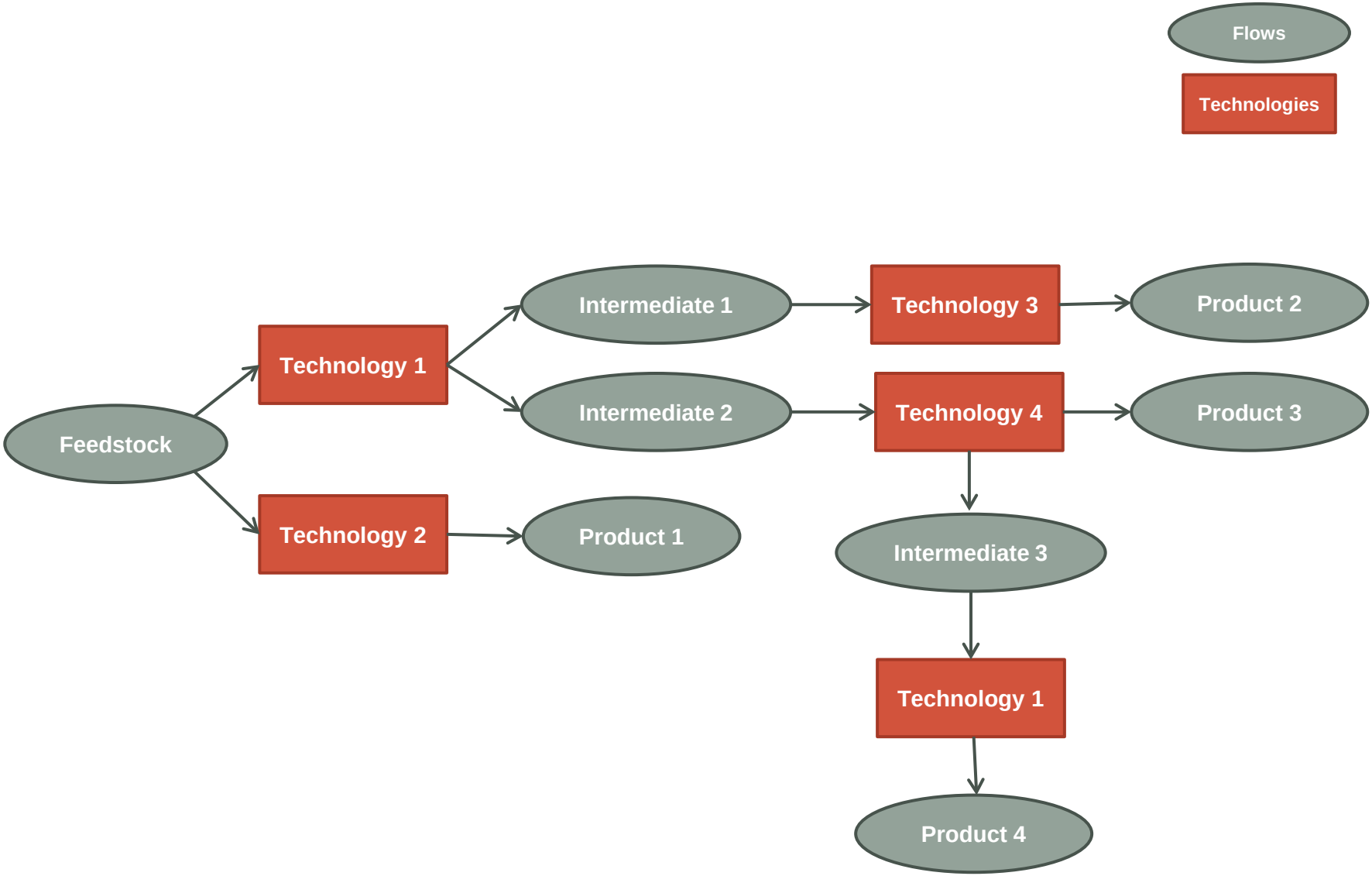
Thomas R. Gruber (1993)

Ontology engineering

- **Classes**: collection of entities that share a common characteristic
- **Instances / Individuals**: entities that belong to a particular class
- **Relations / Object properties** (Domain, Range)
- **Data properties**: e.g. *Age* (Peter,15)

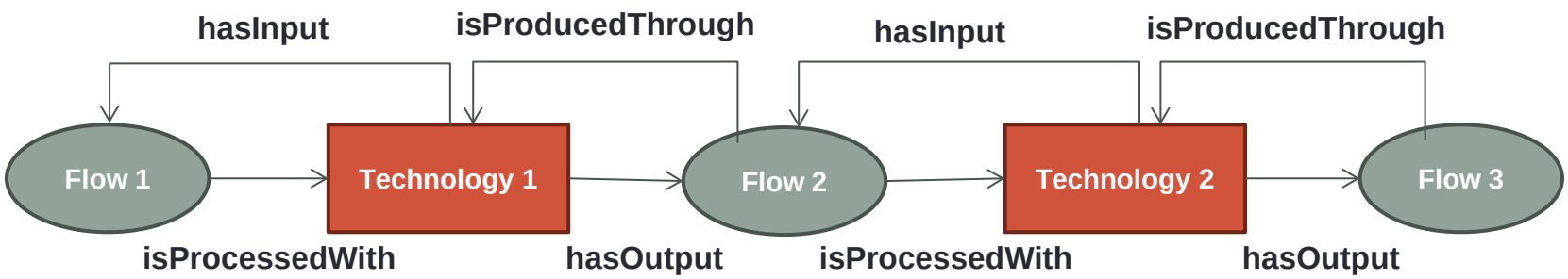


An ontology for biorefineries – Synthesis of paths



An ontology for biorefineries – Object properties

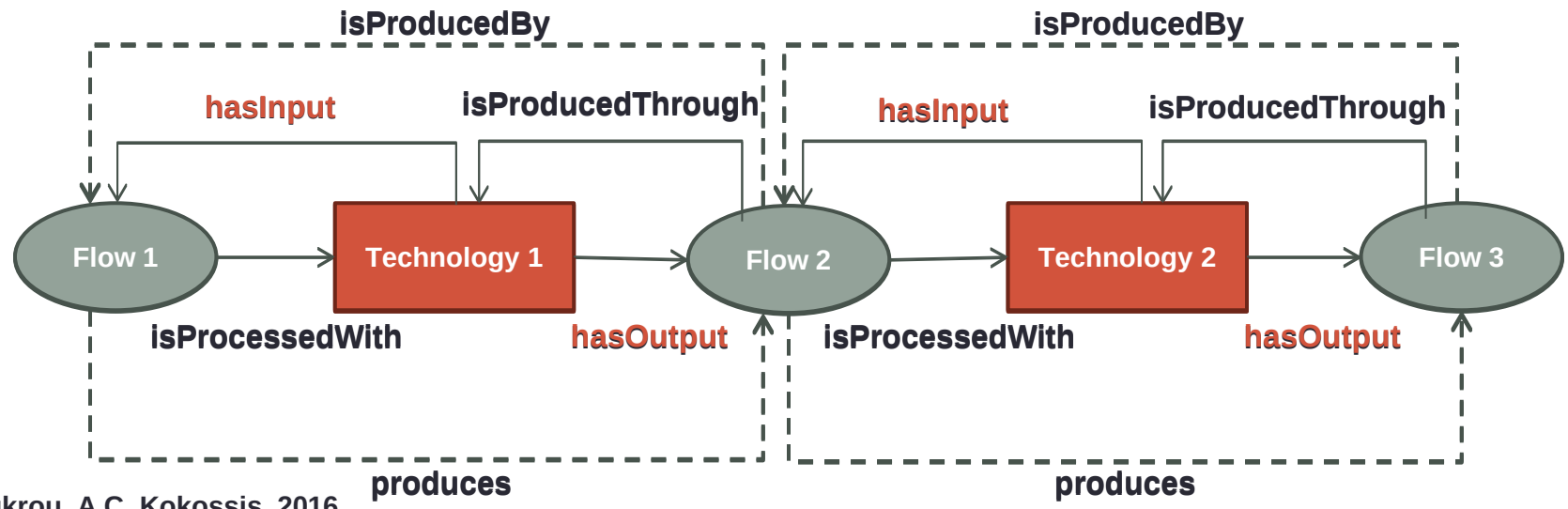
- **hasInput** (Technologies,Flows)
- **IsProcessedWith** (Flows,Technologies) ***Inverse of*** **hasInput**
  - Automatic connection by reasoner
- **hasOutput** (Technologies,Flows)
- **IsProducedThrough** (Flows,Technologies) ***Inverse of*** **hasOutput**
  - Automatic connection by reasoner



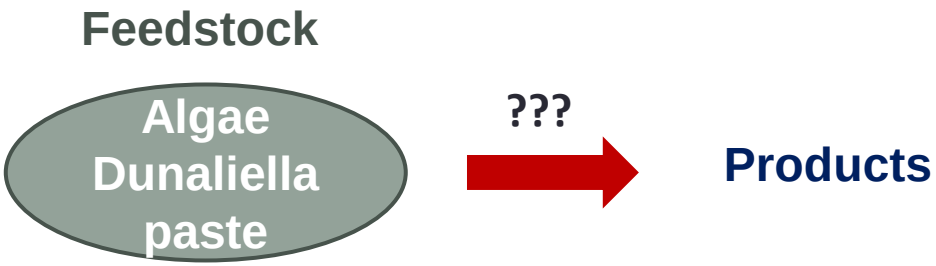
An ontology for biorefineries – Object properties

chain

- **produces** (Flows,Flows) : **IsProcessedWith o hasOutput**
  - Automatic connection by reasoner
- **isProducedBy** (Flows,Flows) ***Inverse of*** **produces**
  - Automatic connection by reasoner



Synthesis of paths / value chains



Synthesis of paths / value chains

Value chain algorithms

Step 1. Select Feedstock F

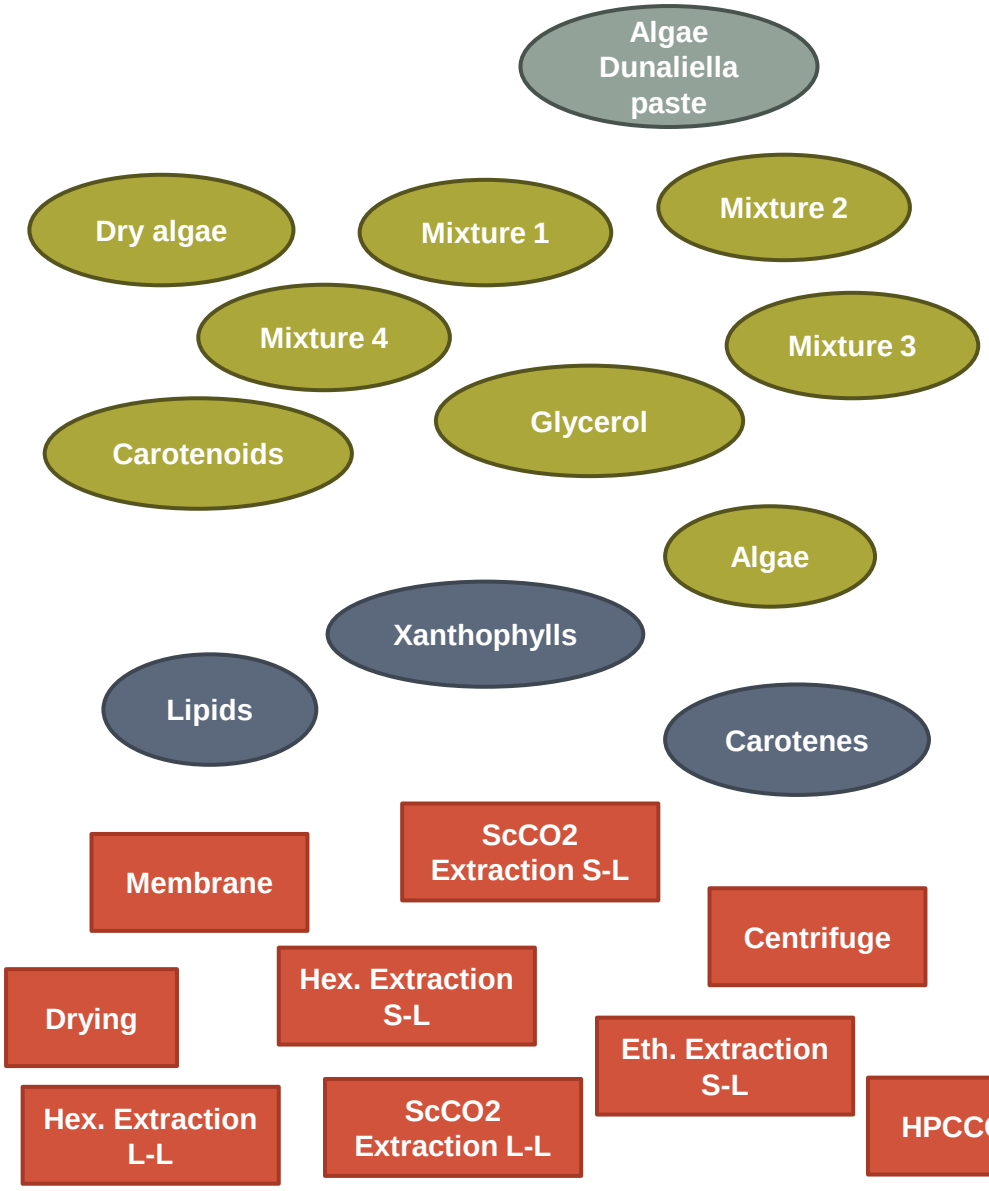
Step 2. Find Intermediates I & Products P where F *produces* I *&* P

Step 3. Find Technologies T where T *hasInput* F *&* T *hasOutput* I *OR* P

Step 4. Find Technologies T where T *hasInput* I *&* T *hasOutput* I *OR* P

Step 5. Find Technologies T where T *hasInput* P *&* T *hasOutput* I *OR* P

Step 6. Transfer pathways to graph

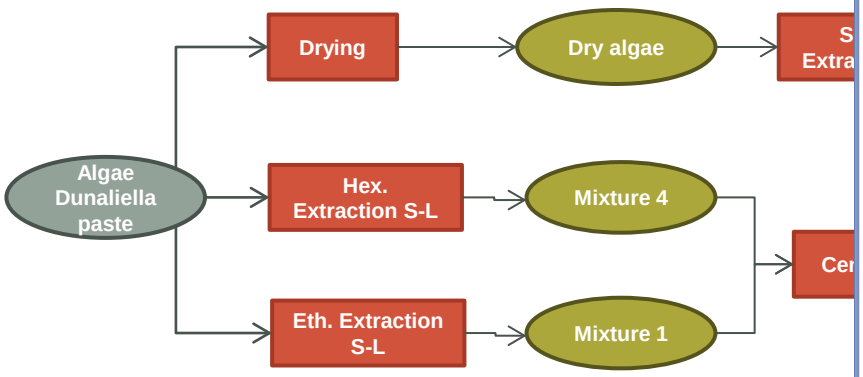


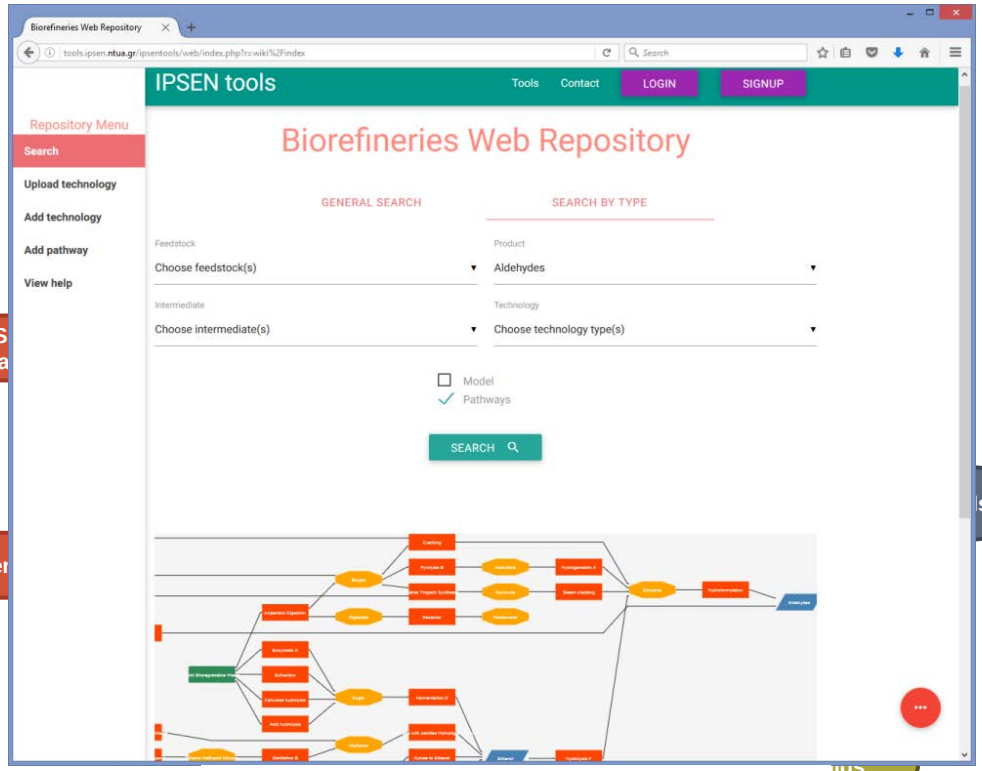
E. Siougkrou, F. Lykokanellos, A.C. Kokossis, in preparation

Synthesis of paths / value chains

Value chain algorithms

Step 6. Transfer pathways to graph

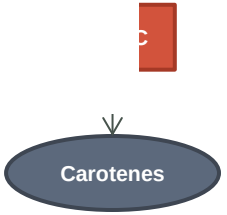




Link with optimisation formulation to find the optimal path

Available on the web

URI: tools.ipsen.ntua.gr



Summary - Perspectives

- Ontology engineering is a promising alternative to mathematical programming
  - User-friendly applications
- Ontology engineering offers the technology for high-throughput organisation of large data and complex relations
- Many possible applications, e.g. industrial symbiosis, integration of technologies/industries (1G + 2G biorefineries), metabolic pathways

Acknowledgements

Collaborators

- Prof. Claire Adjiman
- Prof. Amparo Galindo
- Prof. Antonis Kokossis
- Dr. Heiko Struebing
- Dr. Zara Ganase
- Mr. Fil Lykokanellos
- Ms. Foteini Barla
- Mr. Stefan Oberhaimer

Thank you!

Funding

- EPSRC, Molecular Systems Engineering grant (EP/E016340)
- Marie Curie European Research Program, RENESENG (FP-607415)