Table of Contents

1. Laboratorium Języków Assemblerowych 2
1.1. Ćwiczenie 1 2
1.1.1. Przygotowanie: Środowisko sprzętowe i programistyczne 2
1.1.1.1. Instalacja Visual Studio (Community Edition 2019) 3
1.1.2. Wymagania 4
1.1.3. Wstęp 4
1.1.4. Powtórka z teorii: Funkcje i ich wywoływanie z poziomu assemblera 5
1.1.4.1. Standardy wywoływania funkcji 5
1.1.5. Praktyka: Tworzenie projektu C++ i konsersja do ASM 7
1.1.5.1. Tworzenie pliku z kodem źródłowym 7
1.1.5.2. Konwersja projektu z C++ na ASM 8
1.1.5.3. Utworzenie punktu startowego 10
1.1.6. Zadanie do wykonania 12
1.1.7. Debuggowanie i uruchamianie 12
1.1.8. Metoda weryfikacji wyników 15
1.1.9. Pomocne zasoby 15

1. Laboratorium Języków Assemblerowych

1.1. Ćwiczenie 1

Laboratorium języków assemblerowych, ćwiczenie 1

1.1.1. Przygotowanie: Środowisko sprzętowe i programistyczne

Do wykonania ćwiczenia niezbędny jest komputer PC w postaci maszyny fizycznej lub wirtualnej z systemem Windows 7, 8, 8.1 lub 10, z zainstalowanym i skonfigurowanym środowiskiem programistycznym Microsoft Visual Studio 2019 w wrsji co najmniej Community Edition [1)]. Do pracy w laboratorium nada się również wersja VS Professional lub Enterprise. Proszę natomiast nie pomylić z Visual Studio Code, to zupełnie inny produkt, który nie nadaje się wprost do wykorzystania w naszym laboratorium do realizacji zadań, w szczególności kolejnych ćwiczeń.



Jednym z najprostszych rozwiązań do przygotowania środowiska laboratoryjnego jest wykorzystanie gotowej maszyny wirtualnej, która jest dostępna na stronach Microsoft. Można ją wykorzystać w celach testowych, w okresie nie dłuższym niż 30-90dni (zależnie od daty pobrania i wersji). Dostępne są maszyny wirtualne dla różnych platform wirtualizacyjnych.

W przypadku chęci skorzystania z gotowej maszyny, obraz można pobrać m.in. z poniższego linku:

- Opcja 1: Windows 10 + Visual Studio 2019 Enterprise [2)];
- Opcja 2: Windows 10, a następnie własnoręczna instalacja Visual Studio 2019 Community[3)];

Jako środowisko witualizacyjne można wykorzystać Oracle Virtual Box[4)], dostępne praktycznie na każdą platformę desktopową. W przypadku instalacji Linuxowych typu headless, można wykorzystać KVM[5)].



Maszyny wirtualne z pewnością będą wymagały zainstalowania aktualizacji Windows i aplikacji, w tym kilkukrotnych restartów, zatem instalację należy wykonać zawczasu i doprowadzić ją do stanu stabilnego PRZED zajęciami.

Proszę rownież pamiętać o zainstalowaniu dodatków specyficznych dla platformy wirutalizacyjnej (np, Oracle Virtual Box Add-ons), co może pociągać za sobą konieczności w zmianie konfiguracji maszyny wirtualnej (np. konieczność dołożenia napędu wirtualnego w celu zainstalowania ww dodatków). Sugeruję rownież zwiększenie pamięci RAM do co najmniej 8GB (o ile to możliwe), liczby dostępnych wątków procesora (przynajmniej 2) oraz zmianę wirtualnej karty graficznej na zgodną z akceleracją sprzętową DirectX/Direct 3D (np. w Oracle Vitual Box jest to karta VBoxSVGA).



Niezależnie od przyjętego modelu należy pamiętać, że zajęcia wymagają 64-bitowego

systemu operacynego.

1.1.1.1. Instalacja Visual Studio (Community Edition 2019)

Instalator Visual Studio Communicy Edition 2019 pozwala na zainstalowanie wielu różnych komponentów. Minimalny zestaw komponentów został przedstawiony na poniższych rysunkach 1, 2, 3:

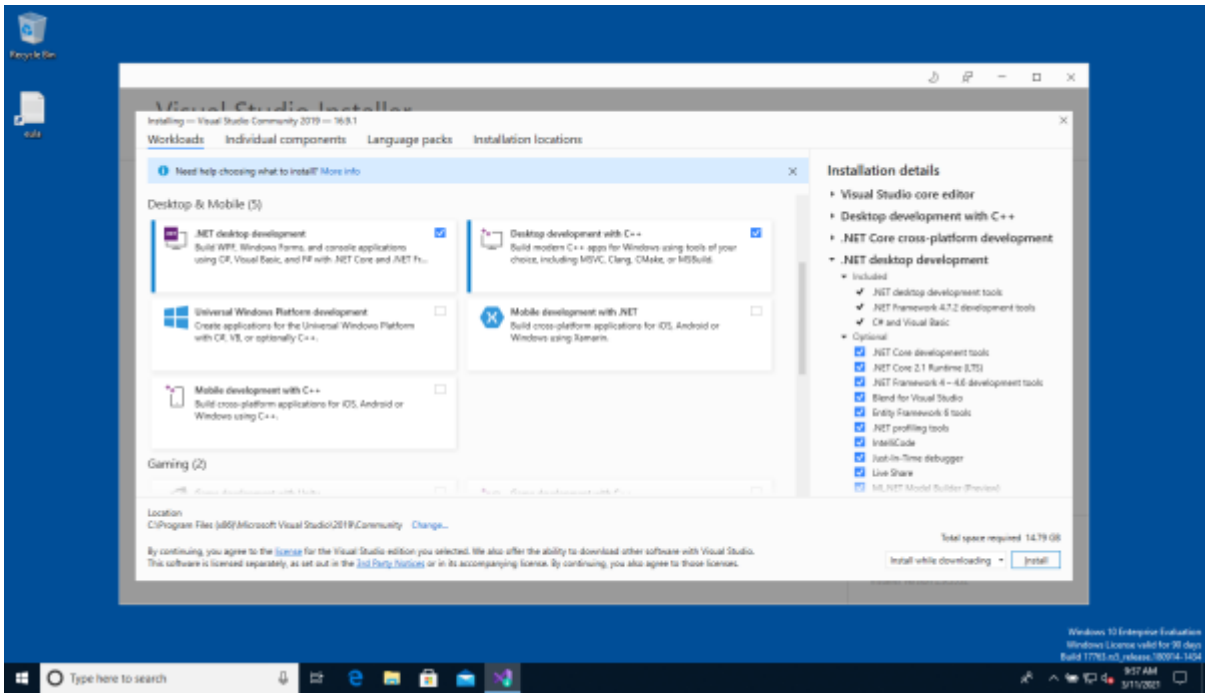


Fig. 1: Opcje instalacyjne część 1

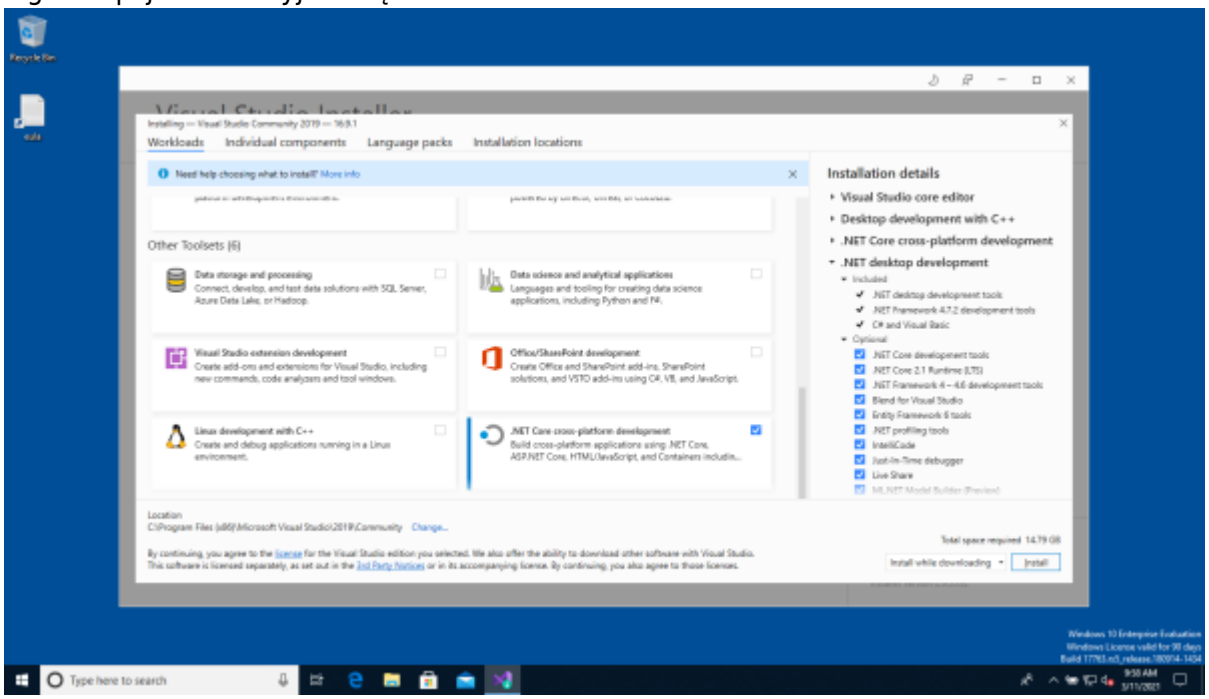


Fig. 2: Opcje instalacyjne część 2

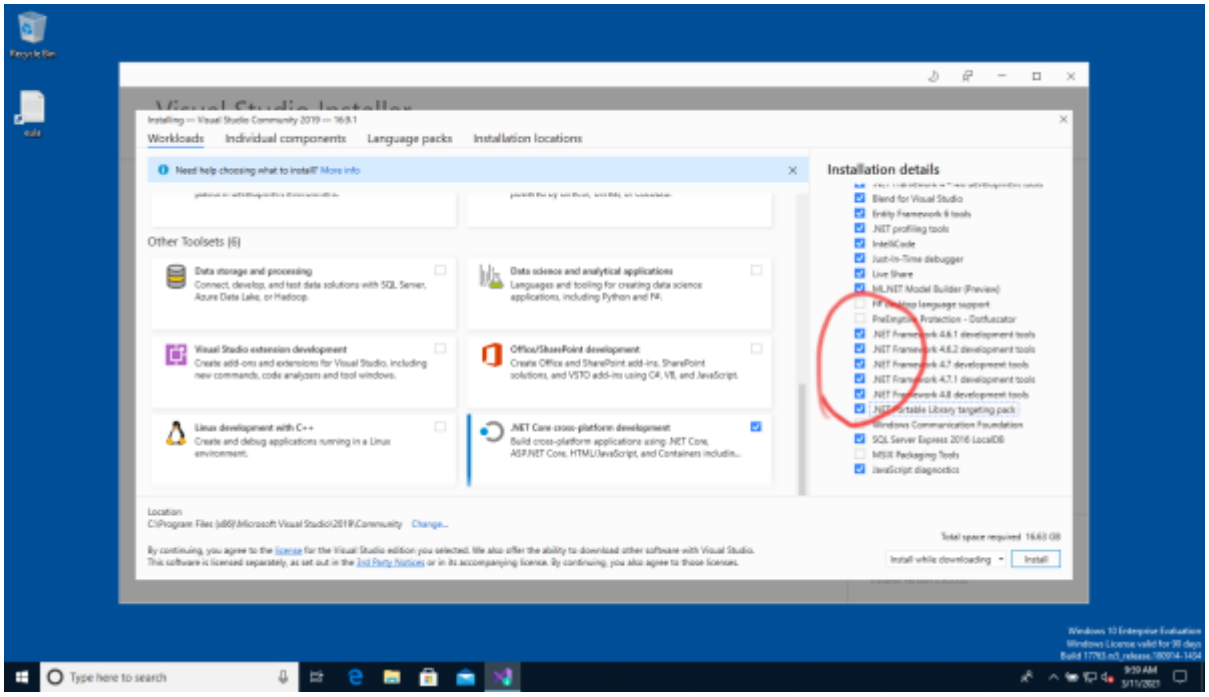


Fig. 3: Opcje instalacyjne część 3 - uwaga - frameworki .NET po prawej stronie trzeba zaznaczyć ręcznie!

1.1.2. Wymagania

Studenci muszą rozumieć sposób tworzenia i działania aplikacji typu “Command Line”, podstawy języka assembler, C++ i C#

1.1.3. Wstęp

W ramach tego ćwiczenia stworzysz w czystym assemblerze dla x64 (procesory Intel/AMD, 64 bity) aplikację konsolową, która będzie prowadziła interakcję z systemem operacyjnym Windows 10 (7, 8, 8.1) wywołując jego funkcje poprzez interfejs Windows API.

Jest to przykład technologiczny, ponieważ tworzenie w ten sposób skomplikowanych aplikacji i złożonych interfejsów użytkownika jest ... bezsensowne 😬, choć możliwe. Natomiast zdedycowanie przyda się przy robieniu prostych usług obliczeniowych, narzędzi, etc., które wymagają zwartości kodu i wykonywania zadań z maksymalną prędkością.
Zwróć uwagę, że podane przykłady nie ograniczają się do funkcji interfejsu użytkownika, a obejmują całe Windows API. **A ze względu na rozmiar aplikacji, całość niejednokronie mieści się w pamięci L1 procesora, zatem wykonuje się niespotykanie szybko.**

Ponieważ w Visual Studio nie ma gotowego modelu projektu na potrzeby aplikacji w czystym assemblerze, wykorzystamy model aplikacji konsolowej w języku C++ i przezbroimy go na potrzeby kompilacji zamiast kodu w języku C++, kodu w czystym assemblerze.

Takie podejście wymaga ręcznego przekonfigurowania kilku pramatetrów we właściwościach projekt, a w szczególności:

- zmiany kompilatora na `masm`;
- zmiany docelowej platformy na x64;
- podania nazwy głównej funkcji jako punktu wejścia ([Entry Point]) tak, aby sytem operacyjny

wiedział, którą funkcję ma zawołać po uruchomieniu aplikacji - w wielu językach i środowiskach programowania jest to domyślnie funkcja `main()`, natomiast w przypadku assemblera, nie ma takiej reguły i trzeba podać samemu, o którą funkcję chodzi;
- uzupełnienia brakujących bibliotek: od wersji 2019 środowiska (w szczególności pod Windows 10) cześć standardowych funkcji dostępnych uprzednio domyślnie w bibliotekach linkowanych do każdego projektu, aktualnie zostało wyłączonych i trzeba je jawnie wskazać; uwaga - dokładny skład bibliotek jest zależny oczywiście od tego, co jest wykorzystywane (wywoływane) w kodzie źródłowym - w środowisku produkcyjnym należy zoptymalizować kod poprzez usunięcie zbędnych bibliotek;



Uwaga! Ponieważ wykorzystujemy pliki `asm`, które natywnie nie są rozpoznawane przez Visual Studio jako pliki z kodem źródłowym, nie są one automatycznie zapisywane w momencie wywołania kompilacji, jak ma to miejsce w przypadku plików `.cpp`, `.cs` itp. Co prawda istnieje możliwość nauczenia VS 2019, aby plik był uznawany za źródłowy, niemniej jednak wykracza to poza niniejszą instrukcję. Proszę zatem pamiętać za każdym razem kliknąć zapis do pliku (Ctrl+s) przed kompilacją.

1.1.4. Powtórka z teorii: Funkcje i ich wywoływanie z poziomu assemblera

W przypadku pisania kodu w języku assembler, każdy bajt musi być przydzielony, zadeklarowany i użytkowany w odpowiedni sposób, szczególnie w kontekście wywoływania funkcji Windows API. Ponieważ pracujemy na systemie 64.-bitowym, konieczne jest stosowane reguł wywołania funkcji zgodnie z kanonem.

1.1.4.1. Standardy wywoływania funkcji

Parametry wywołania funkcji przekazujemy poprzez rejestry, o ile ich liczba nie przekracza czterech. W przeciwnym wypadu przekazujemy je przez stos (pozostałe). Natomiast obowiązkowo należy zaalokować na stosie przestrzeń na argumenty nawet ... jeśli przekazujemy je przez rejestry 🙂[6].

1.1.4.1.1. Przekazanie argumentów całkowitych

W przypadku, gdy do funkcji przekazujemy argumenty numeryczne całkowite, w szczególności np. adresy i wskaźniki, należy użyć w kolejności następujących rejestrów:

1. argument 1 → RCX,
2. argument 2 → RDX,
3. argument 3 → R8,
4. argument 4 → R9,
5. każdy kolejny argument umieszczamy na stosie z zachowaniem szesnastobitowego wyrównania danych (z przyczyn wydajnościowych).

Przykładowy kod w C++:

```
func1(int a, int b, int c, int d, int e, int f);
```

```
// a w RCX, b w RDX, c w R8, d w R9, f, a następnie e, są odkładane na
stosie
```

1.1.4.1.2. Przekazywanie argumentów zmiennoprzecinkowych

W przypadku przekazywania argumentów zmiennoprzecinkowych, kolejne wartości ładujemy do rejestrów XMM:

1. argument 1 → XMM0,
2. argument 2 → XMM1,
3. argument 3 → XMM2,
4. argument 4 → XMM3,
5. analogicznie jak dla stałoprzecinkowych.

Przykładowy kod w C++:

```
func2(float a, double b, float c, double d, float e, float f);
// a jest XMM0, b w XMM1, c w XMM2, d w XMM3, f a następnie e są odkładane
na stosie
```

1.1.4.1.3. Przekazywanie argumentów mieszanych

W przypadku mieszania argumentów zmiennoprzecinkowych z dziesiętnymi, stosuje się system pozycyjny w zależności od typu, a zatem jeśli przekazujemy np, int1, float1, float2, int2, to wartości należy umieścić w rejestrach (w kolejności):

1. int1 → RCX,
2. float1 → XMM1,
3. float2 → XMM2,
4. int2 → R9,
5. piąty i następny argument poprzez stos.

Przykładowy kod w C++:

```
func3(int a, float b, float c, int d, int e, float f);
// a w RCX, b w XMM1, c w XMM2, d w R9, f, a następnie e, poprzez stos
```



Rejestry po wywołaniu funkcji ulegają zniszczeniu (zmianie wartości), zatem nie należy oczekiwać, że po wywołaniu funkcji, rejestry RCX, RDX, R8, R9, XMM0-XMM3 będą miały wartości takie, jak przed wywołaniem funkcji!

1.1.4.1.4. Zwrot wartości z funkcji

Wartość całkowita (w tym wskaźnik) zwracana jest w rejestrze RCX.

Wartości zmiennoprzecinkowa zwracana jest w rejestrze XMM0.

1.1.5. Praktyka: Tworzenie projektu C++ i konsersja do ASM

Stwórz nowy projekt: aplikacja typu command-line (konsolowa) dla języka C++, natywnego (nie C++ .NET) rysunek 4

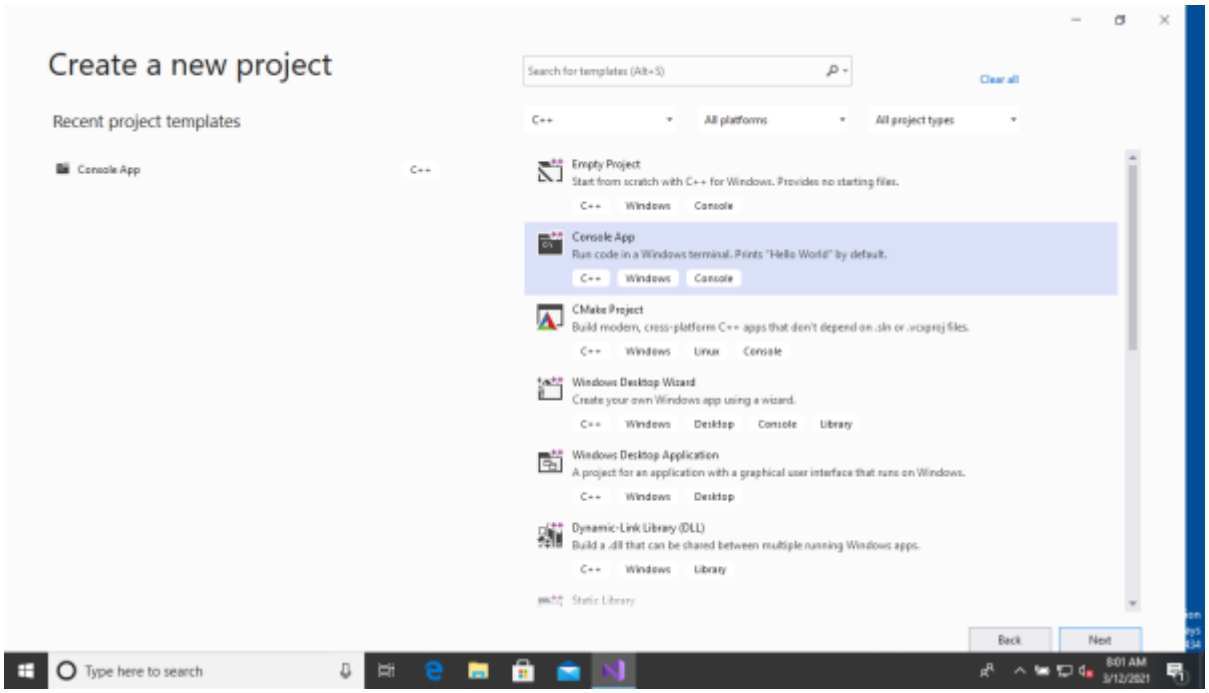


Fig. 4: Tworzenie projektu typu C++, kod natywny

1.1.5.1. Tworzenie pliku z kodem źródłowym

Usuń wygenerowany plik .cpp (z poziomu VS) a następnie dodaj nowy plik .asm (można to zrobić klikając w projekt (nie w solucję!) i wybierając dodanie nowego elementu rysunek 5, a następnie wybrać plik typu tekstowego i zmienić jego nazwę oraz rozszerzenie rysunek 6:

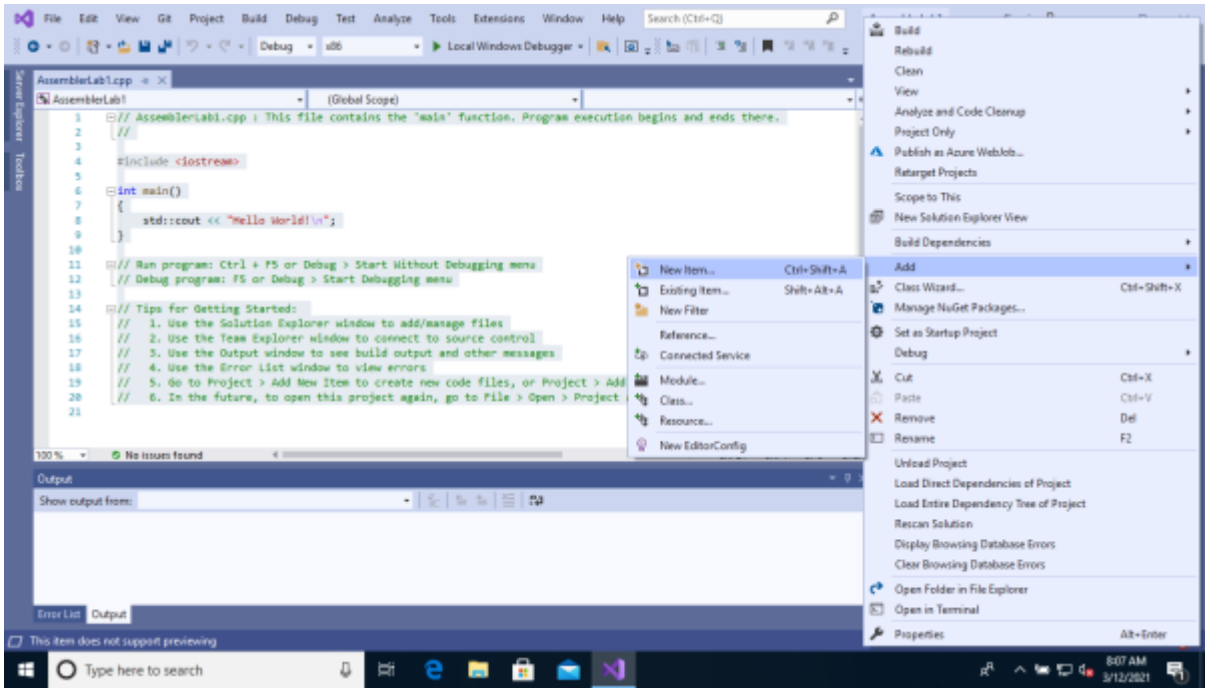


Fig. 5: Dodanie nowego elementu

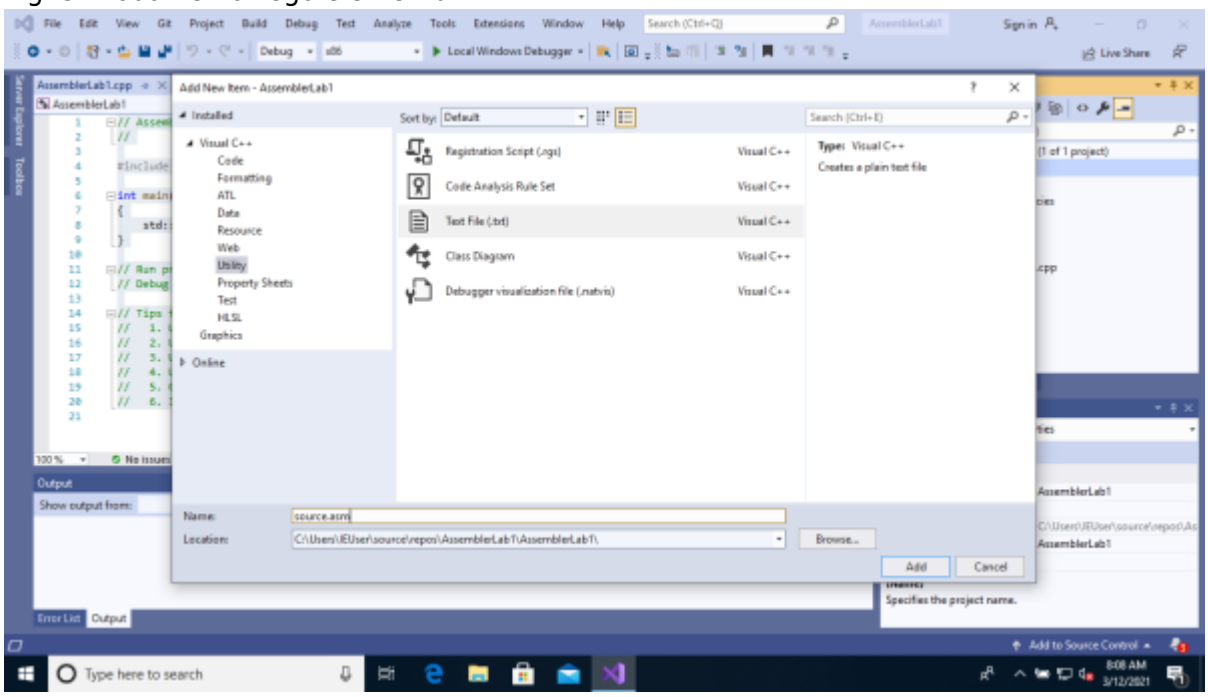


Fig. 6: Dodanie nowego elementu - plik asm

1.1.5.2. Konwersja projektu z C++ na ASM

W celu konwersji należy poinformować VS, że pliki źródłowe będą kompilowane za pomocą MASM a nie kompilatora C/C++. W tym celu klikamy prawym przyciskiem na projekt i wybieramy opcję dostosowania budowania projektu rysunek 7. W ustawieniach należy wybrać `masm(.targets, .props)` rysunek 8.

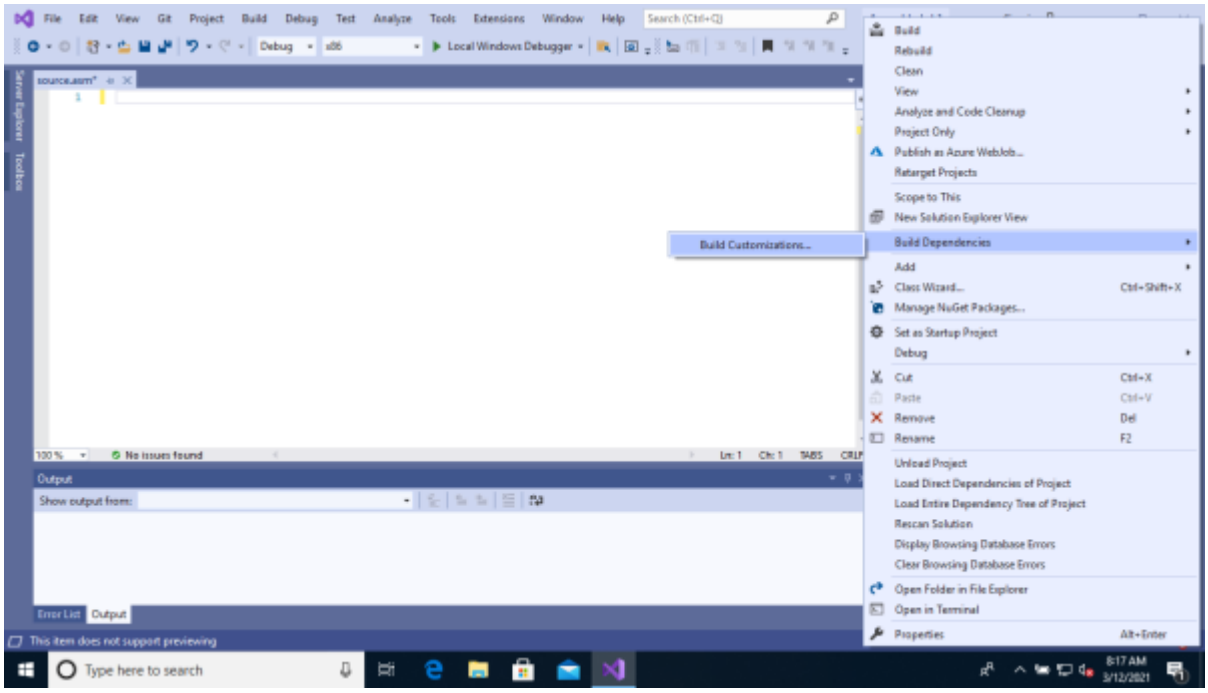


Fig. 7: Wybór menu dostosowanie budowania projektu

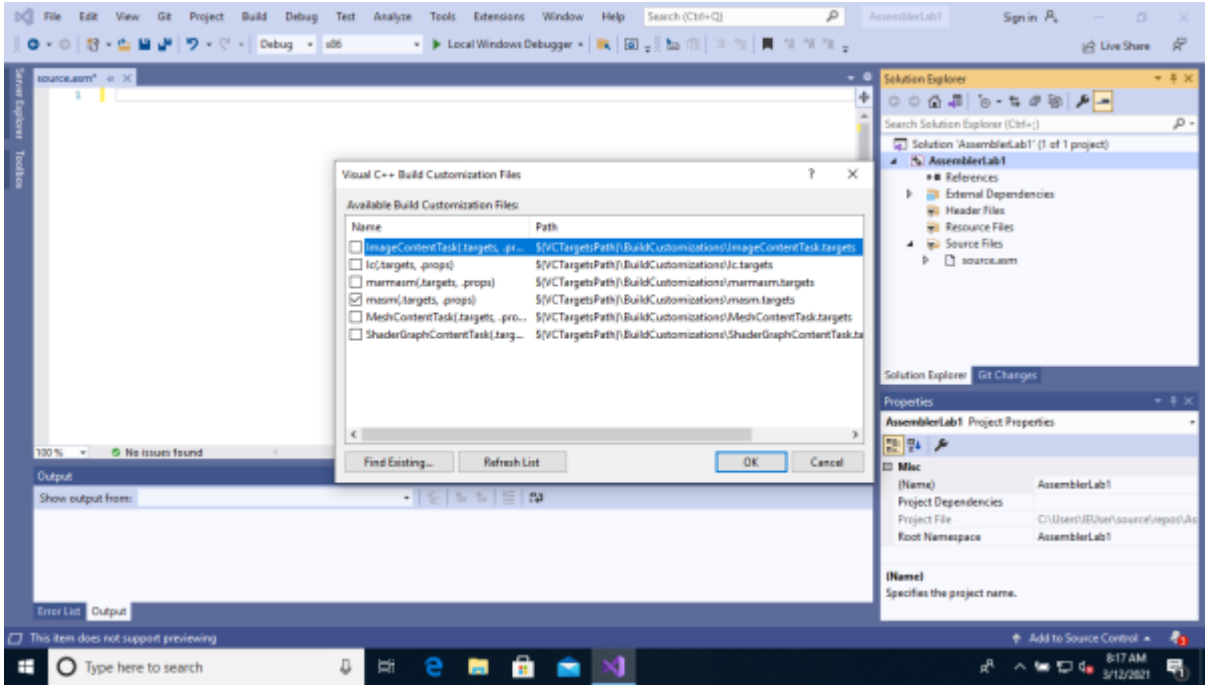


Fig. 8: Wybór kompilatora MASM

Następnie należy zmienić rodzaj budowanej aplikacji z x86 (aplikacja 32-bitowa) na x64 (aplikacja 64-bitowa): menu Build→Configuration Manager rysunek 9.

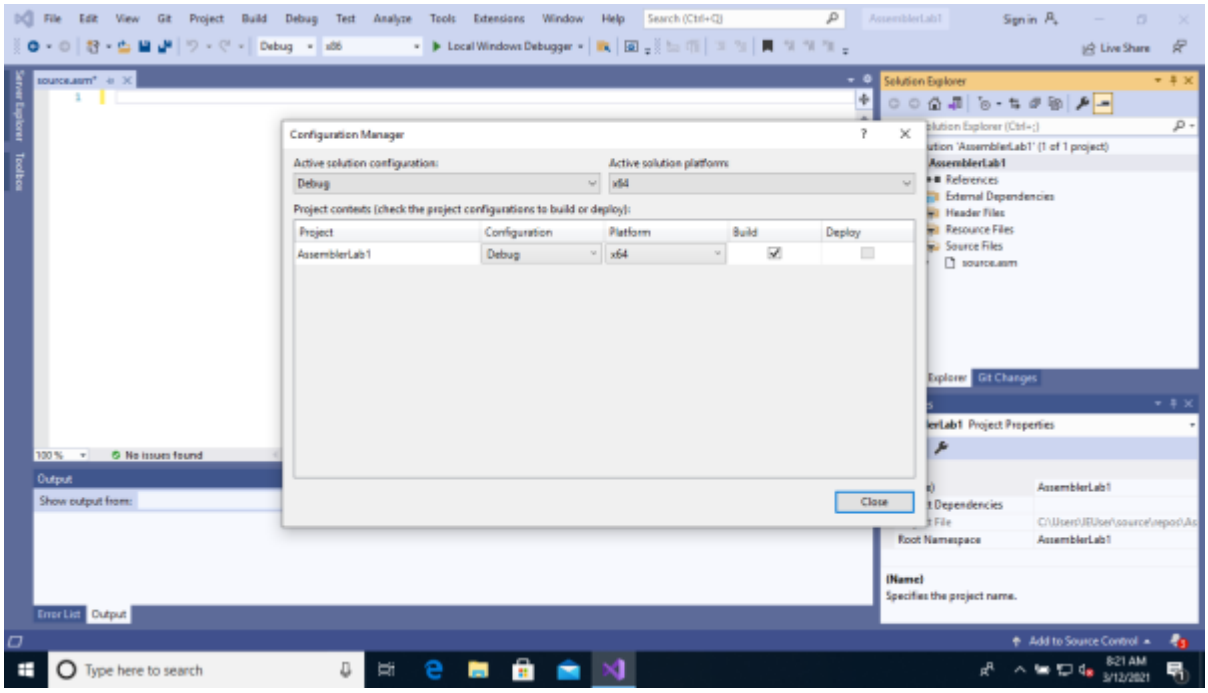


Fig. 9: Konfiguracja środowiska - budowa wersji 64-bitowej

1.1.5.3. Utworzenie punktu startowego

W tym kroku poinformujemy linker, jaki punkt startowy ma wpisać do pliku wykonywanego. W językach wysokiego poziomu jest to domyślnie funkcja `main()`. W przypadku assemblera nie ma takiego założenia, a zatem kod źródłowy musi być zgodny z ustawioną tutaj wartością.



Podana do linkera nazwa funkcji punktu wejścia do kodu musi zgadzać się z kodem źródłowym. W przypadku rozbieżności, program będzie działał niepoprawnie, nie uruchomi się lub rzuci wyjątkiem przy starcie!

W podanym przykładzie, nazwa funkcji to `mojaFunkcjaAsm`. Aby dostać się do okna konfiguracyjnego należy w drzewie projektu wskazać projekt (nie solucję!), a następnie prawym przyciskiem myszy kilknąć w projekt, wybrać pozycję `Properties`, lub na skróty [Alt+Enter]. Pojawi się okienko jak na rysunku 10. Następnie w sekcji `Linker` należy wybrać opcję `Advanced` i w polu `Entry Point` wpisać nazwę funkcji.

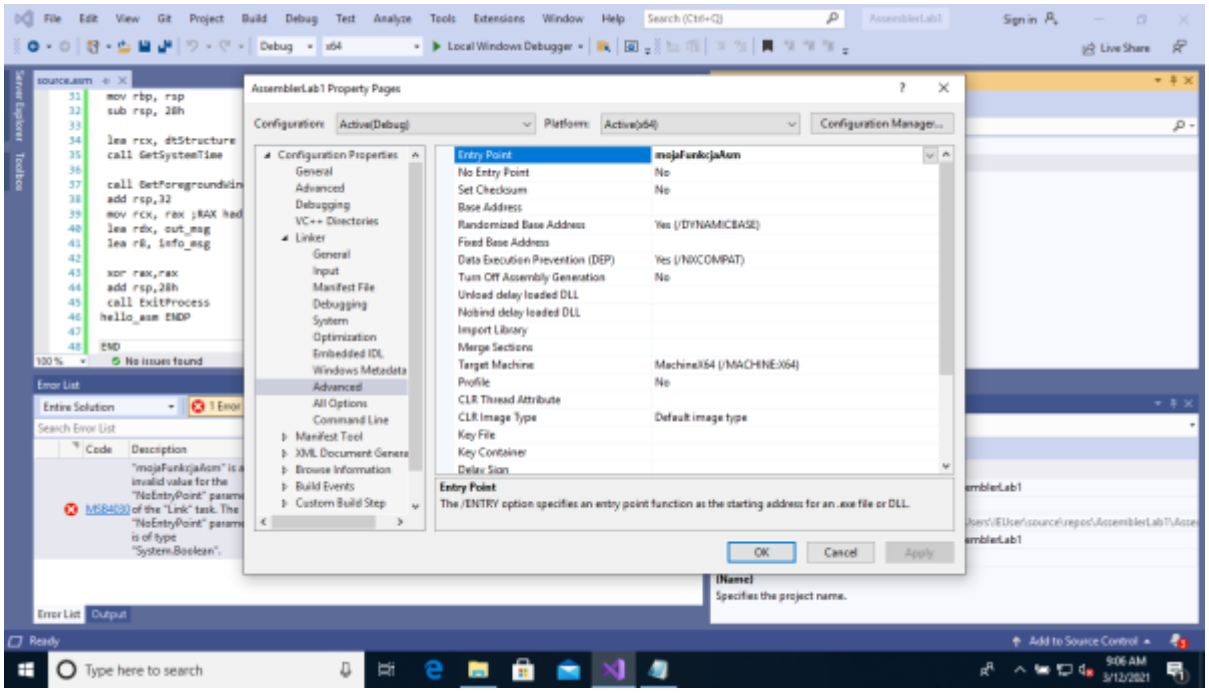


Fig. 10: Konfiguracja punktu wejścia do programu

W tej sekcji należy również dołożyć brakujące biblioteki, z których będziemy korzystać. Uwaga - w zależności od wersji Windows i bieżącego poziomu aktualizacji, ten krok może nie być potrzebny, ale nie zaszkodzi. W szczególności w tych bibliotekach znajdują się finkcje typu `sprintf`.
Uwaga - podane biblioteki należy dołożyć do tych, które standardowo już tam się znajdują. Usunięcie znajdujących się tam bibliotek (m.in. `kernel32.lib`,`user32.lib`) spowoduje brak możliwości użycia standardowych okienek dialogowych oraz interakcji z Windows API. Aby dołożyć biblioteki należy wskazać w menu `Linker` pozyję `Input`, a następnie wybrać `Additional Dependencies` i opcję `<Edit..>` po rozwinięciu listy wyboru. Następnie w górnym polu wpisać nazwy plików bibliotecznych, które mają być dostępne dodatkowo podczas kompilacji (jedna pozycja w linii). Biblioteki do uzupełnienia to:

- `legacy_stdio_definitions.lib`
- `legacy_stdio_wide_specifiers.lib`

Szczegóły na rysunku 11.

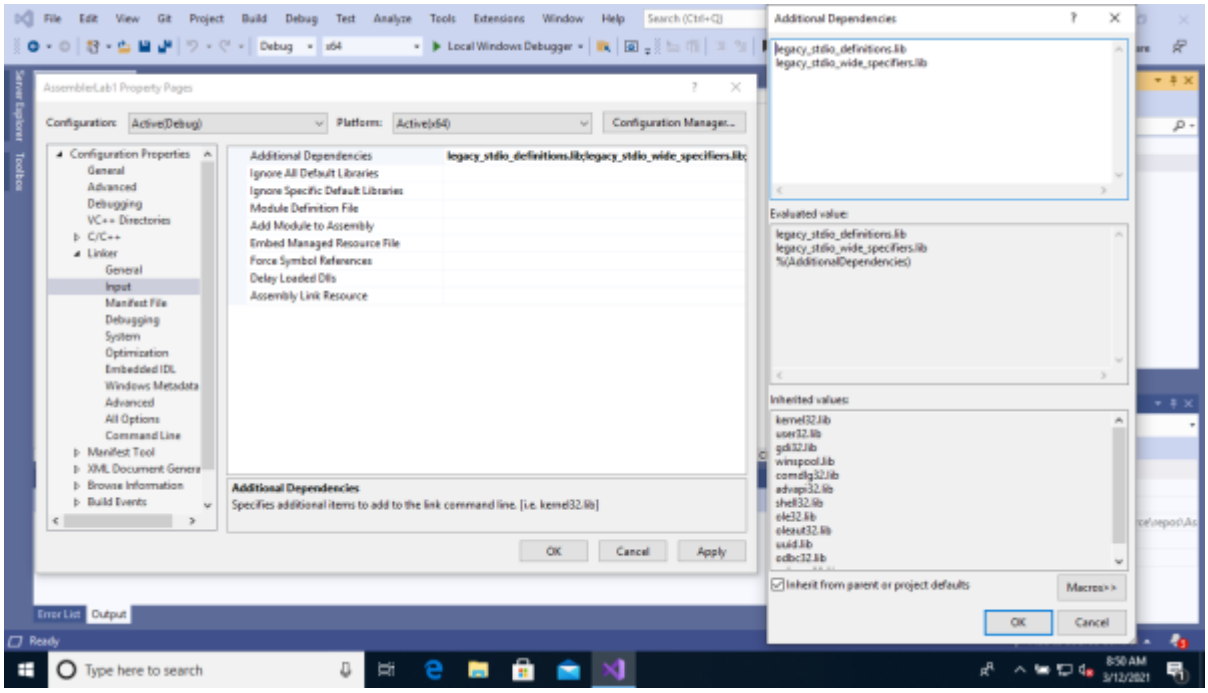


Fig. 11: Konfiguracja bibliotek dodatkowych

1.1.6. Zadanie do wykonania

Zaimplementuj odczyt daty (użyj funkcji `GetSystemTime`(sysinfoapi.h)[7] odczytującej strukturę `_SYSTEMTIME`(minwinbase.h)[8]. Struktura składa się z 8 wartości szesnastobitowych.



Uwaga - nie deklaruj struktury, jako takiej, tylko dla uproszczenia zadeklaruj docelową, statyczną tablicę 8 wartości typu Word (16bit) w segmencie danych.

Skonwertuj rok, miesiąc i dzień do postaci tekstowej ASCII i wyświetl na ekranie za pomocą:

- w linii poleceń (cmd) za pomocą funkcji `WriteConsoleA`;
- w okienku dialogowym, za pomocą funkcji `MessageBoxA`;



Zwróć uwagę na położenie poszczególnych słów w strukturze i odpowiadające im wartości.

Możesz dodatkowo wyświetlić bieżącą godzinę, minuty i sekundy.

1.1.7. Debuggowanie i uruchamianie

Uwagi praktycznie odnośnie debbugowania i podglądania pamięci:

- Pamiętaj, aby na poziomie wywołań funkcji Windows API nigdy nie wykonywać operacji wejścia do wewnątrz ich definicji (Step Into, F11), o ile wcześniej nie zainstalowałeś Windows SDK wraz z definicjami i kodem źródłowym. Zamiast tego należy używać przejściea do następnej instrukcji (Step Over, F10)
- Okno rejestrów domyślnie pokazuje tylko podstawowy zestaw: więcej po kliknięciu prawym

przyciskiem myszy i włączeniu innych zestawów rejestrów
- Okno podglądu pamięci pozwala na zmianę formatowania danych z domyślnego, bajtowego, na wiele innych. Np w kontekście podglądu ww. struktury składającej się z danych 16-bitowych zdecydowanie warto zmienić podgląd na 2-bajtowy. Natomiast do podglądania ciągów znaków ASCII należy wrócić do jednobajtowego.



Aby podejrzeć zawartość pamięci, najprościej jest odnaleźć adres zmiennej w rejestrze, po wykonaniu instrukcji LEA(Load Effective Address). Wartość z okienka podglądu stanu rejestrów należy przekopiować do okienka podglądu stanu pamięci.



Uwaga - powyższą operację należy wykonać pod debuggerem zaraz po wykonaniu instrukcji LEA ponieważ następujący po niej CALL powoduje zmiany zawartości rejestrów i adres zostanie zamazany.

Na rysunku rysunek 12 przedstawiono kolejność działań w celu odczytania (odnalezienia) miejsca w pamięci, gdzie znajduje się zmienna. Breakpoint (pułapkę) należy ustawić bezpośrednio na następnej instrukcji znajdującej sie po LEA, chyba że jest gwarancja, iż instrukcje pomiędzy LEA i breakpointem nie modyfikują danego rejestru. Odczytaną wartość należy przenieść z okienka rejestrów do okienka pamięci. Po wywołaniu funkcji, zawartość rejestru RCX zmienia się, natomiast dane zwracane są do pamięci, jak na rysunku rysunek 13. Zadanie dodatkowe: zamień widoczną na rysunku zawartość pamięci, a dowiesz się, kiedy powstał screenshot 😉.

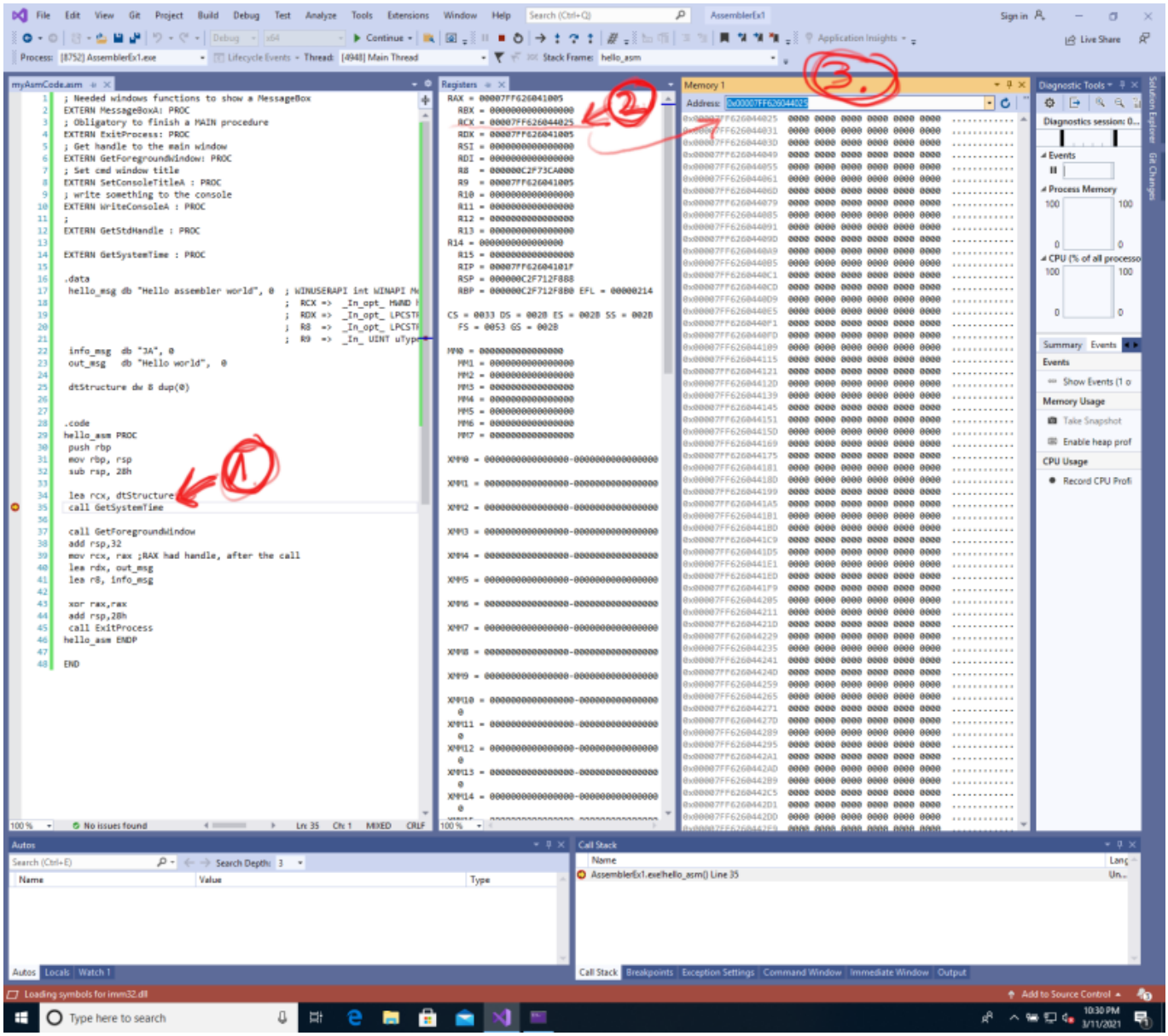


Fig. 12: Wyciąganie adresu dla zmiennej (stan bezpośrednio po wykonaniu LEA)

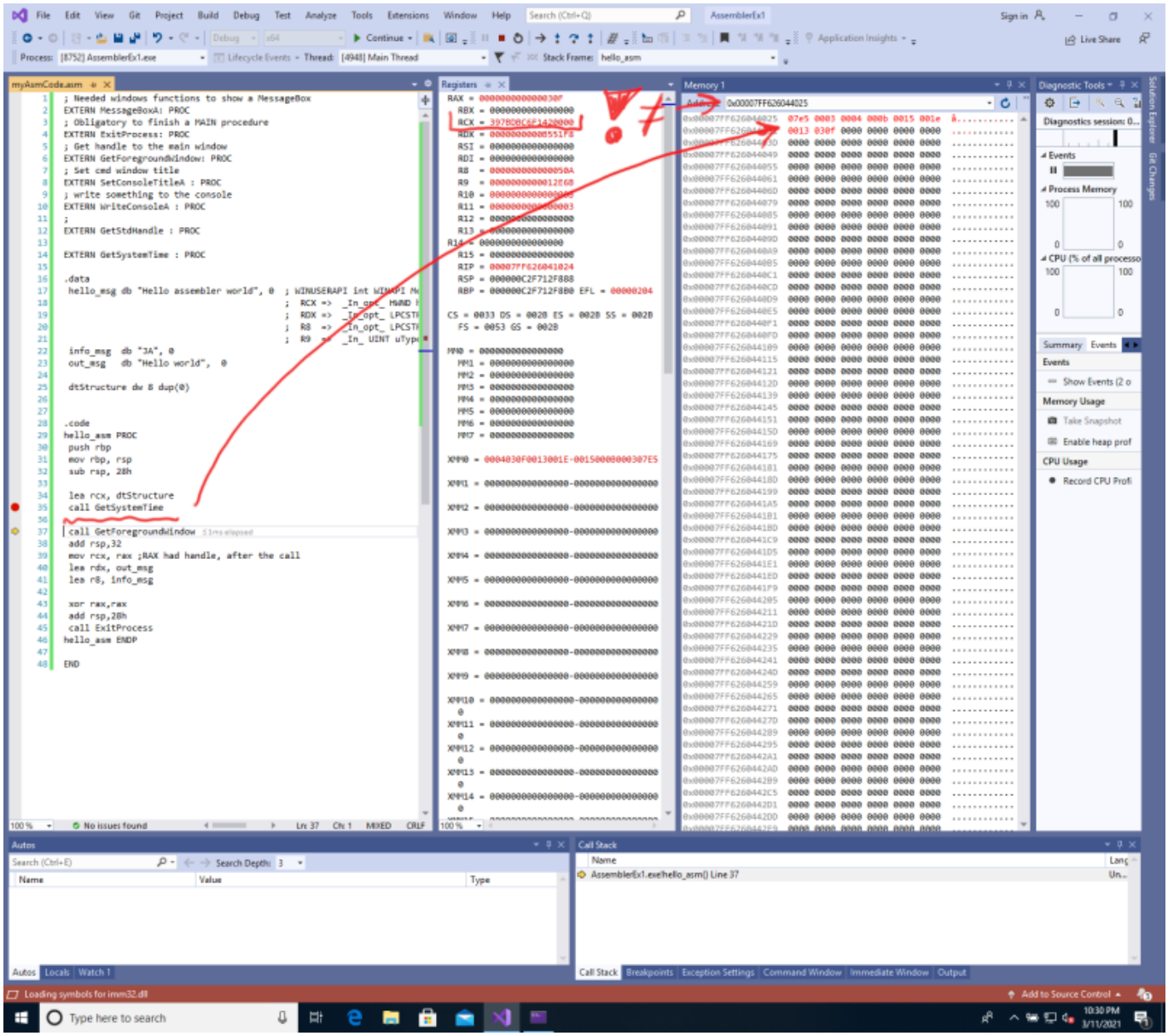


Fig. 13: Obserwacja wyniku w pamięc, po wywołaniu funkcji

1.1.8. Metoda weryfikacji wyników

Teksty oraz okienka pojawiające się na ekranie zgadzają się z algorytmem. Zwróć szczególną uwagę na to, czy nie są generowane losowe ciągi znaków.

1.1.9. Pomocne zasoby

- Visual Studio 2019: https://visualstudio.microsoft.com/vs/
- Przydatne funkcje Windows API: https://docs.microsoft.com/en-us/windows/win32/api/_winmsg/

1)

https://visualstudio.microsoft.com/vs/

2)

https://developer.microsoft.com/en-us/windows/downloads/virtual-machines/

3)

https://developer.microsoft.com/en-us/microsoft-edge/tools/vms/

4)

https://www.virtualbox.org

5)
https://www.linux-kvm.org/page/Main_Page
6)
https://docs.microsoft.com/en-us/cpp/build/x64-calling-convention?view=msvc-160
7)
https://docs.microsoft.com/en-us/windows/win32/api/sysinfoapi/nf-sysinfoapi-getsystemtime
8)
https://docs.microsoft.com/en-us/windows/win32/api/minwinbase/ns-minwinbase-systemtime

From:
https://dokuwiki.zmitac.aei.polsl.pl/ - **RAU6 Dokuwiki**

Permanent link:
https://dokuwiki.zmitac.aei.polsl.pl/pl:edu:zmitac:n1z:ja:lab1?rev=1615565317

Last update: **2021/03/12 17:08**